



TAMPEREEN TEKNILLINEN YLIOPISTO
TAMPERE UNIVERSITY OF TECHNOLOGY

JORI LÄHTEELÄ

MYYN TIENNOSTUKSIEN TARKKUUDEN SIMULOINTI

Diplomityö

Tarkastajat:
Apulaisprofessori Minna Lanz ja
Tohtorikoulutettava Ville Toivonen
Tarkastajat ja aihe hyväksytty
29.11.2017.

TIIVISTELMÄ

JORI LÄHTEELÄ: Myyntiennustuksien tarkkuuden simulointi

Tampereen teknillinen yliopisto

58 sivua

Tammikuu 2018

Automaatiotekniikan koulutusohjelma

Pääaine: Tuotantotekniikka

Tarkastajat: Prof. Minna Lanz ja Ville Toivonen

Avainsanat: Ennustaminen, tarkkuuden mitta

Tässä työssä tutkitaan ennustustarkkuuden mittaamista. Tätä tutkitaan virittämällä Arima X11 -ennustusmenetelmä ennustamaan yhden yrityksen myyntiä. Ennustettavat frekvenssit ovat kuukausi ja kvartaaliennusteet, joita tehdään koko myynnille, neljälle eri yrityksen sisäiselle myyntikanavalle tai maittain. Lisäksi tutkitaan viritykseen käytetyn historiatiedon pituuden vaikutusta saatavaan tarkkuuteen.

Aineistona käytettiin vuosien 2008 ja 2016 välistä ABB Medium Voltage Products myyntihistoriaa. Tarkkuuden mittaamisen käsitteen tarkasteluun käytettyä aineistoa olivat akateemiset tutkimukset, mittaustekniikan opinnot ja aiheesta julkaistu muu kirjallisuus.

Tarkkuusmittarina käytettiin yrityksen sisäistä prosentuaalista mittaria sekä absoluuttista prosentuaalista virhettä. Tarkkuusmittareiden vahvuuksia ja heikkouksia käsitellään teoriaosuudessa. Tarkkuus mitataan simuloimalla kuukausiennustukset 12 kuukautta ja kvartaaliennustukset 8 kvartaalia taaksepäin käytetyn historiatiedon lopusta ja vertaamalla saatuja ennusteita vastaaviin todellisiin myynteihin. Virityspituuden vaikutusta saataviin tarkkuuksiin tutkitaan ajamalla ennustukset kaikilla mahdollisilla virityspituuksilla. Näistä tuloksista nähdään miten saatavat tarkkuudet muuttuvat kullakinkin käytetyllä mittarilla virityspituuden funktiona.

Tutkimuksen johtopäätöksenä todetaan, että pidemmän historiatiedon käytöllä ei välttämättä saavuteta tarkempia ennusteita. Virityshistorian vaikutusta saatavaan tarkkuuteen ei voida yleistää, sillä se riippui paljon ennustettavasta sarjasta ja käytetystä mittarista. Käytetyistä mittareista toinen saattaa parantua toisen heikentyessä, joten mittarin valinnan tärkeys korostuu, mikäli halutaan valita paras ennustemalli tiettyyn tarkoitukseen. Käytetty ennustusmenetelmä täytti kaiken myynnin kuukausi- ja kvartaalimyyntien ennustamisen odotukset 100%, kanavien kuukausiennustamisen 88% ja maiden ennustukset 55% tarkkuudella.

ABSTRACT

JORI LÄHTEELÄ: Simulation of sales forecasting accuracies

Tampere University of Technology

Master of Science, 58 pages

January 2018

Master's Degree Programme in Automation Technology

Major: Production Engineering

Examiner: Prof. Minna Lanz ja Ville Toivonen

Keywords: Forecasting, Accuracy analysis

In this study the accuracy of sales forecasting is being measured and analysed. To accomplish this Arima X11-forecasting method is used to produce month and quartile forecasts for the sales of a company. Forecasted series are all sales, four internal sales channels and country based series. In addition the dependency of sales history length used for tuning the forecasting model and the accuracy of results is studied.

The sales history from ABB Medium Voltage Products between years 2008 and 2016 was used as a source for the study. As a source for the theoretical part of the study, academic articles, theoretical literacy about the theory of metrics and other relevant printed material for the topic were used.

For the means of measuring accuracy, company's percentage meter and mean absolute percentage error is used. Caveats for metrics are specified in the theory section. Forecast accuracy is measured by simulating one year backwards from the end of available sales history. This is done for month forecasts and two years backwards for quartal forecasts. Measurement is done by comparing actual and simulated forecasts. The dependancy between the length of the tuning model and accuracy is evaluated by simulating all possible tuning lengths for all needed forecasts.

As a conclusion it is stated that longer sales history used for tuning the model does not necessarily produce better forecasting accuracy. Forecasting accuracy depends greatly on forecasted time series and the selected accuracy metrics. It is possible that two of the used metrics were showing contradictory results for selecting the best length for used history. For this reason it is essentially important to choose the used metrics carefully in order to choose the most accurate forecasting methods. The forecasting method (used in this study) fulfilled expectancies with 100%. Expectancies for month forecasts of the channels were fulfilled 88% and for countries 55%.

ALKUSANAT

Työssä tutkitaan ennustustarkkuuden mittaamista, jota varten on viritetty ennustusjärjestelmä tuottamaan kuukausi- ja kvartaaliennusteita syötetystä historiatiedosta.

Työ valaisi ennustamiseen liittyvää maailmaa ja siihen liittyviä ongelmia laajasti. Kiitän mielenkiintoisen aiheen mahdollistamisesta Mika Kukkolaa, Lasse Alostaa ja Timo Eskolaa, jotka ohjasivat työtäni ABB Oy:n puolesta. Yliopiston tarkastajille Minna Lanzille sekä Ville Toivoselle annan ison kiitoksen pitkäjänteisestä ohjaamisesta, kannustamisesta ja korjausehdotuksista.

Lopuksi kiitän kavereitani, perhettäni, tuttujani ja erityisesti vaimoani Minna Lähteelää, joka on pitkäjänteisesti mahdollistanut minulle aikaa tutkimuksen tekemiseen.

Tampere, 17. tammikuuta 2018

Jori Lähteelä

SISÄLLYS

1. Johdanto	1
2. Mallinnus	3
2.1 Dynaaminen, staattinen vai stationaarinen systeemi	4
2.2 Muuttujien relaatiot	5
2.3 Datan esikäsittely	7
2.4 Mallin tyyppi	9
2.4.1 X11-komponentteihin jako	13
2.5 Ylivirittäminen	15
3. Tarkkuuden mittaus	17
3.1 Koejärjestely	18
3.2 Arvosarjan osittelu	19
3.3 Suoritusmittari	20
3.4 Virheanalyysi	23
4. Simulointijärjestelmä	26
4.1 Tekninen toteutus	26
4.2 Ennustusten muodostaminen	28
4.2.1 Ennustettavat sarjat	28
4.2.2 Virityspituus	29
4.2.3 Ennustus ja tulostenkäsittely	29
4.2.4 Loppuvuoden ennusteen toteutus	30
4.3 Tarkkuusmittarien valinta	30
4.4 Tarkkuuksien simulointi ja esitys	32
5. Analyysi	36
5.1 Kuukausiennusteet	36

5.2 Kvartaaliennusteet	40
6. Tulokset	45
7. Yhteenveto	50
Lähteet	54

1. JOHDANTO

Tutkimus sai alkunsa tutkiessani työssäni ABB Oy:llä myyjien tekemien myyntiennusteiden tarkkuuksia. Pohdin olisiko mahdollista löytää jokin matemaattinen ennustemetodi, jolla saavutettaisiin vähintään nykyiset ennustetarkkuudet. Tästä syntyi idea diplomityöhöni rakentaa simulointijärjestelmä ennustemallien vertailuun, jotta useiden ennustemallien toimivuutta ja toimivuutta voitaisiin verrata keskenään.

Ennustuksien tarkkuuden analysoiminen mahdollistaa sopivimman mallin valinnan lisäksi niiden kehittämisen ja siten ennustuksien tarkkuuden paranemisen. Ennustamisesta voidaan saada hyötyä, kuten esimerkiksi Xiande Zhao, Jinxing Xie ja R. S. M. Lau totesivat tuotantoketjun eri osapuolten hyötynneen suuresti sen ohjaukseen otettaessa käyttöön asiakastilauksiin perustuva ennustus[1]. M. Babai, M. Ali, J.Boylan ja A.Syntetos puolestaan mallinsivat ja ennustivat kaksivaiheista tuotantoketjua [2] ja totesivat ennustamisen parantavan tehokkuutta noin 40%. Ennustamisen hyödyllisyyteen päätyivät myös myös P. Danese ja M. Kalchschmidt tutkiessaan ennustamisen vaikutusta operatiiviseen tehokkuuteen[3].

Mallien toimiessa keskenään eri tavoin myös niiden vahvuusalueiden oletetaan olevan poikkeavia. Tästä syystä päädyttiin mittaamaan ennustetarkkuuksia usealla eri mittarilla, jotta jokaiseen tilanteeseen voitaisiin valita sopivin ennustemetodi. Tilastollisessa ennustamisessa ennustemallin viritykseen käytetään historiallisia myyntitietoja. Tämän historiatiedon käytöstä nousee esiin ongelma kannattaako käyttää kaikkea saatavilla olevaa historiatietoa vai saataisiinko tarkempia ennusteita jollakin lyhemmällä virityshistorialla. Tutkimusongelmaksi asetettiin Arima x11 ennustemallin tuottamien ennusteiden tuottamien myyntiennusteiden tarkkuuden mittaamisen lisäksi tarkkuuden analysoiminen virityspituuden funktiona.

Tutkimusaineisto koostuu ABB Medium Voltage Products -yksikön myynnistä jaoteltuna useaan eri ennustettavaan arvosarjaan. Jaottelu tehdään jakamalla ennustettava myynti maiden, yrityksen neljän sisäisen myyntikanavan ja koko yksikön myyntiin (Kuva 1.1). Näistä muodostuvia myyntejä ennustetaan kuukausittain

ja kvartaaleittain, joista saatavia ennustetarkkuuksia analysoidaan. Tutkimukses-

Taulukko 1.1 *Analysoitavat ennustetarkkuudet voidaan jakaa kuuteen osa-alueeseen.*

Data			
Kaikki	Kanavat	Maat	Ennusteaika
1.	2.	3.	Kuukausi
4.	5.	6.	Kvartaali

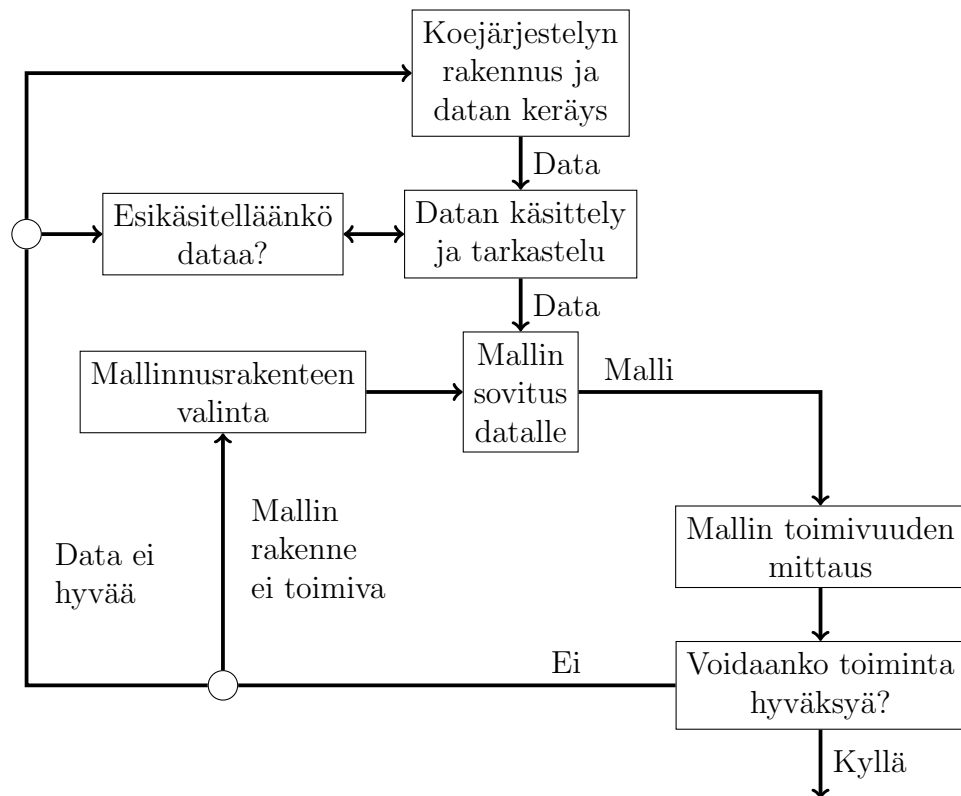
sa käytetty myyntitieto rajautuu vain yhden yrityksen yhteen yksikköön, joten saatavat tulokset eivät ole välttämättä suoraan yleistettävissä muihin tilanteisiin ja käyttökohteisiin. Valitussa ennustemetodissa sarjoja ennustetaan sarjakohtaisesti, joten niiden mahdolliset ristikkäisvaikutukset jätetään huomioimatta. Maiden lukumäärästä johtuen niihin laskettavat ennusteet mahdollistavat tilastollisen tarkkuusanalyysin.

Tutkimusmenetelmä on tarvittavien ennusteiden simulointi eri virityspituuksilla, joiden perusteella voidaan laskea tarkkuudet valituilla tarkkuusmittareilla. Tarkkuuden mittaamisen ajanjaksoksi on kuukausiennusteille valittu 12 kuukautta ja kvartaaliennusteille 8 kvartaalia, joiden ajalta saatavia tarkkuuksia tutkitaan. Tarkkuusmittareina on käytetty yrityksen sisäisen tarkkuusmittarin ja euromääräisen virheen lisäksi käytetyimpiin mittareihin lukeutuvaa absoluuttista prosentuaalista keskivirhettä.

Työn teoriaosassa esitellään tarkkuusmittarien ja mittaustapojen vaikutusta saataviin ennustetarkkuuksiin. Soveltavassa osassa esitellään miten ennustetarkkuutta mittaava simulointijärjestelmä on rakennettu mittaamaan Arima x11-ennustusmetodin kuukausi- ja kvartaaliennustetarkkuuksia. Analyysissä kootaan yhteen saadut ennustetarkkuudet. Yhteenveto esittelee tutkimuksessa esille tulleita käytetyn ennustusmenetelmän vahvuuksia ja heikkouksia.

2. MALLINNUS

Malli voi olla matemaattinen tai verbaalinen kuvaus systeemin toiminnasta. Sitä käytetään työkaluna vastaamaan systeemiä koskeviin kysymyksiin, jotta ei tarvitsisi tehdä reaali maailman kokeita. Matemaattisen mallin avulla on mahdollista simuloida (Lat. *simulare* = teeskennellä) systeemin toimintaa, jotta voidaan arvioida miten se olisi voinut käyttäytyä todellisessa tilanteessa. [4]



Kuva 2.1 Mallinnus on iteroiva prosessi, jota jatketaan kunnes mallin toiminta on hyväksyttävissä. Muokattu lähteestä [4].

Mallinnus on iteroiva prosessi, jonka aikana mallintaja tarkentaa mallia sen tuottamien tuloksien perusteella kunnes tarkkuus on hyväksyttävissä (Kuva 2.1). Ma-

temaattisen mallin ei ajatella koskaan kuvaavan täydellisesti kohteena olevan systeemin luonnetta vaan sen oletetaan olevan riittävän käyttökelpoinen tietyssä tilanteessa, jota tarkennetaan tarpeen mukaan. Mallin objektiivinen käyttökelpoisuuden mittaaminen (Kuva 2.1) on siten oleellinen osa mallinnusta, jotta saataisiin selvä kuva mallin vahvuuksista ja kehittämisen kohteista.

Tämä luku käsittelee mallinnuksessa esiintyviä valintoja, joita mallintajan on tehtävä prosessin eri iteraatiokierrosten aikana. Luku koostuu systeemin luonteen valinnasta, millaiset muuttujat pitäisi löytää, millaista datan esikäsittely voi olla, miten käytettävä malli voi valikoitua ja mitä on ylivirittäminen.

2.1 Dynaaminen, staattinen vai stationaarinen systeemi

Ennen mallin rakentamista valitaan mallinnettavan systeemin luonne. Systeemit voidaan jakaa luonteen mukaan dynaamisiin, stationaarisin ja staattisiin systeemeihin.

Dynaamisen prosessin odotusarvo μ_t muuttuu ajan mukaan, joten sen sanotaan olevan ajasta riippuva systeemi [5]. Sarja luokitellaan dynaamiseksi mikäli sarjassa on nähtävissä trendejä, kausivaihteluita tai muita ajasta riippuvia säännöllisyyksiä (Kuva 2.2). Usein dynaamiseen systeemiin vaikuttaa ajasta riippuvan komponentin lisäksi satunnaisuutta ε , jolloin saadaan

$$y_{t \text{ dynamic}} = \mu_t + \varepsilon, \quad \varepsilon \sim iid(0, \sigma^2) \quad (2.1)$$

. Dynaamisessa sarjassa ei ole pysyvää keskiarvoa μ , jota suuntaan järjestelmä koko ajan pyrkisi stationaarisen sarjan tapaan [6]. Tämän takia dynaamisia sarjoja voidaan stationarisoida ensimmäisen tai toisen asteen erosarjan avulla [5]

$$y'_t = y_t - y_{t-1} \quad (2.2)$$

$$y''_t = y'_t - y'_{t-1} = y_t - 2y_{t-1} + y_{t-2} \quad (2.3)$$

Stationaarisen prosessin odotusarvoon μ eikä jakaumaan vaikuta mittauksen aloitushetki t_1 tai sen kesto $T_{mittaus} = t_2 - t_1$ [7] [8]. Aikasarja luokitellaan stationaariseksi, mikäli siinä ei ole tunnistettavaa jaksonpituutta T (Kuva 2.2) [5].

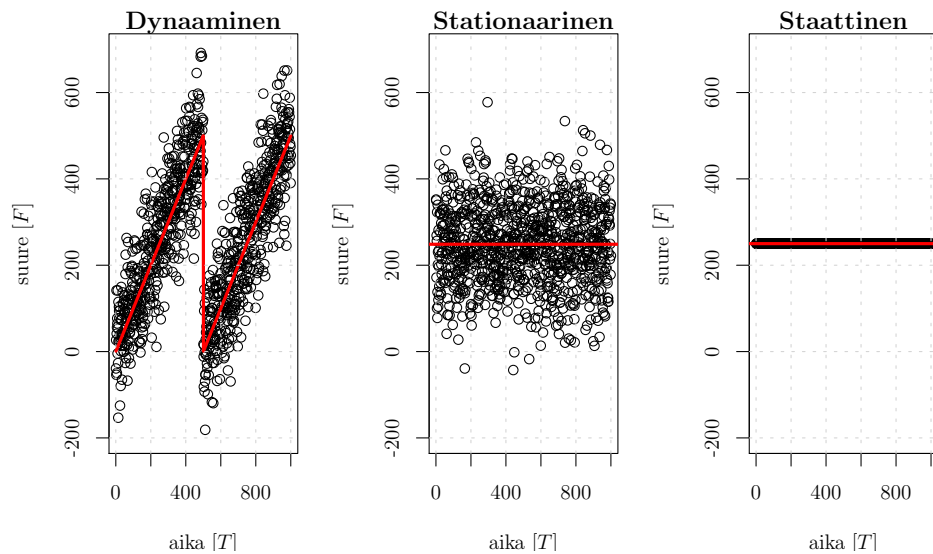
$$y_{t \text{ stationary}} = \mu + \varepsilon, \quad \varepsilon \sim iid(0, \sigma^2) \quad (2.4)$$

Staattisen prosessin odotusarvo μ ei muutu ajan mukaan. Lisäksi sen toiminnassa

ei ole satunnaisuutta ε , joten jokainen ajanhetki systeemin ulostulo $y_{t \text{ static}}$ vastaa aikaisempaa ajanhetkeä (Kuva 2.2).

$$y_{t \text{ static}} = \mu, \text{ joten} \quad (2.5)$$

$$y_{t \text{ static}} = y_{t-1 \text{ static}} = \mu \quad (2.6)$$



Kuva 2.2 Dynaamisen prosessin odotusarvo on riippuvainen ajasta. Stationaarisessa prosessissa ei ole tunnistettavaa jaksonpituutta ja sen odotusarvo tai jakauma ei riipu mittaushetkestä tai sen pituudesta. Staattinen systeemi ei muutu ajan muuttuessa.

Rajanveto erityyppisten prosessien välillä riippuu usein tarkasteltavasta ajasta ja tarkkuudesta. Esimerkiksi rautakappaletta mitattaessa työntömitalla -prosessi voidaan tulkita staattiseksi mikäli mittausvirheellä ei ole suurta merkitystä, stationaariseksi jos mittausvirheen satunnaisuudella on oleellinen merkitys ja dynaamiseksi mikäli tarkastelujakson aikana kappaleen mitat muuttuvat esimerkiksi ruostumisen tai lämpötilan vaihtelun seurauksena merkittävästi. Systeemin luonteen valinta on oleellinen osa mallinnusta sillä se rajaa iteroitavat mallit ja käyttökelpoisen toimialueen.

2.2 Muuttujien relaatiot

Kahden muuttujan välillä sanotaan olevan relaatio mikäli toisen tietäminen helpottaa toisen tietämistä. Relaatiot voidaan jakaa tilastolliseen relaatioon ja toi-

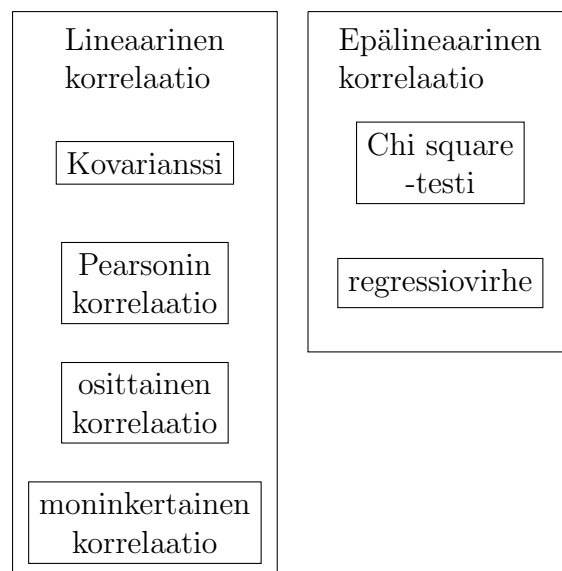
minnalliseen. Toiminnallisen relaation avulla saadaan ensimmäisestä toinen ilman epävarmuutta. Tilastollisella relaatiolla voidaan arvioida ensimmäisestä toisen muuttujan tilastollista todennäköisyyttä. [9]

Korrelaatio on säännöllisyyttä tutkittavien muuttujien käyttäytymisessä. Korrelaatio muuttujien X ja Y välillä ei tarkoita aina kausaalisuutta vaan voi johtua neljästä eri vaihtoehdosta tai niiden yhdistelmistä [10]:

1. Sattumasta (ei kausaalisuutta)
2. X aiheuttaa Y :n
3. Y aiheuttaa X :n
4. Z aiheuttaa X :n ja Y :n

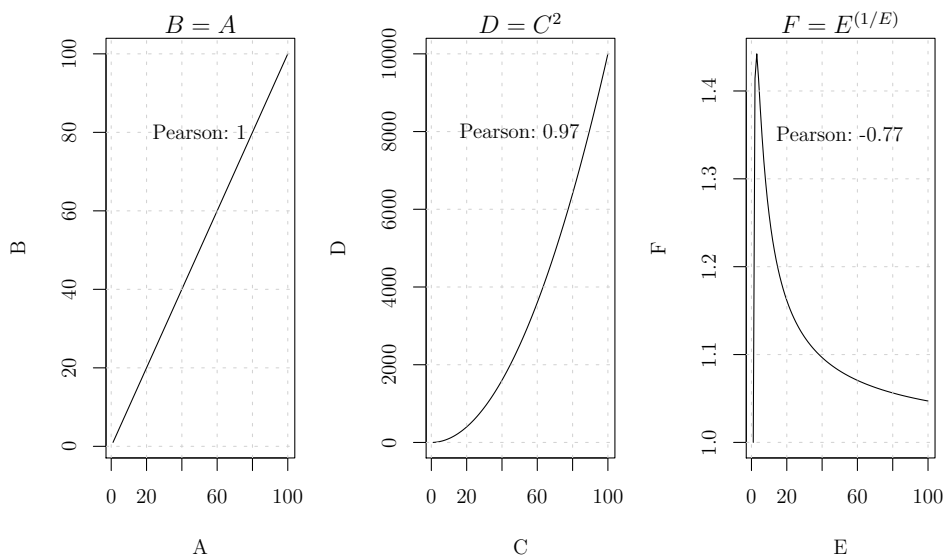
Mikäli suhteessa on selvää säännöllisyyttä piirrettäessä muuttujat toisiaan vastaan (Kuva 2.4), tehdään usein oletus ja oletetaan muuttujien olevan riippuvia toisistaan [9].

Piirtämisen lisäksi muuttujien välisiä suhteita voidaan arvioida numeerisesti erilaisilla korrelaatiomittareilla. Nämä mittarit voidaan jakaa lineaarisiin ja epälineaarisiin (Kuva 2.3), joista jokainen mittari mittaa tietyn tyyppistä ja muotoista relaatiota.



Kuva 2.3 Korrelaatiot voidaan jakaa lineaarisiin ja epälineaarisiin. Muokattu lähteestä [10].

Esimerkki muuttujien toiminnallisten relaatioiden piirtämisestä lisättynä lineaarisella Pearsonin korrelaatiomittarilla on kuvassa 2.4. Tässä jokainen muuttujapari on kausaalinen ja kausaalisuuden tyyppi on ilmoitettu kuvan yläpuolella. Pearsonin korrelaatiomittari voi saada arvoja välillä $[-1, 1]$. Positiivinen luku kertoo positiivisesta suhteesta ja negatiivinen negatiivisesta suhteesta. Luvun itseisarvo kertoo kuinka hyvin suora on sovitettavissa mittauspisteisiin. Mittari saa arvon 0 mikäli siihen ei voida sovittaa suoraa ja muuttujien suhde on täysin satunnainen. [9]



Kuva 2.4 Kolmen eri muuttujaparin väliset suhteet piirrettynä. Pearsonin lineaarinen korrelaatiokerroin kuvaa suhteen lineaarisuutta. Kaikki relaatiot ovat toiminnallisia, joten toinen tiedettäessä toinen voidaan laskea ilman epävarmuutta.

Kuvasta 2.4 nähdään että pelkkä numeerinen korrelaatiokerroin, kuten Pearsonin lineaarinen korrelaatiokerroin, ei ole aina hyvä mittari kausaalisuuden määrittämiseen. Relaation tyyppi ja muodon analysointi on tärkeää oikean mallinnustavan valinnassa. Mikäli sisääntuloiksi valituilla muuttujilla ei ole kausaalista relaatiota ulostuloon, on myös mallin lopullinen tarkkuus huono riippumatta valitusta matemaattisesta mallista.

2.3 Datan esikäsittely

”Roskaa sisään, roskaa ulos” on data-analyysin perussääntö [11]. Mikäli käytettävän datan keräystä ja siinä tapahtuvia virheitä ei tunneta ei voida myöskään luottaa

niistä tehtäviin päätelmiin. A.J. Jaffe ja Herbert F. Spirer ovat jakaneet datan yleisimmät käyttövirheet kolmeen osaan [11]:

1. Virheet ja puutteet
2. Systemaattinen virhe
3. Huonot määrittelyt

Dataa ei voida aina varmistaa. Näistä muodostuvaa virhettä on analysoijan mahdollonta arvioida. Esimerkkejä tällaisesta irrallisesta tiedosta on aborttien määrä kolmannessa maailmassa, nääntymiskuolet kaukana sivistyksen ulkopuolella, alkuasukkaiden määrä viidakossa ja laittomien avaruusolioiden määrä Yhdysvalloissa. Joissain tapauksissa vaikka data voitaisiin varmistaa, se on esitetty väärin ja analysoija ei ole tietoinen siitä. Virheet datan esityksessä voivat tapahtua tahallisesti tai tahattomasti. Mittausvirheet ovat yleisiä toimialasta riippumatta. [11]

Dataan voi syntyä systemaattista virhettä (bias) johtuen keräystavasta. Tällaisia voivat olla johdattelevat tai epämääräiset kysymykset tutkimuksissa. Epämääräisessä kysymyksessä voidaan käyttää monitulkintaista termiä, kuten 'roskaruoka', sitä tarkemmin määrittelemättä [11]. Systemaattista virhettä voi syntyä myös teknisissä mittauksissa vaikka mittalaitteen toleranssista tai inhimillisestä lukuvirheestä. Mikäli mittalaite ajautuu hitaasti pois toleranssista on siitä aiheutuneita virhemittauksia vaikea korjata jälkeenpäin ilman uusintamittauksia. Virhearvion tekemistä vaikeuttaa historiatiedon sisältäessä satunnaisesti tällaista korjaamatonta dataa. Esimerkiksi taloutta mallinnettaessa voidaan pohtia pitäisikö sisääntuleva data korjata inflaatiolla, koska käytettävä mittayksikkö on muuttunut ja miten kyseinen inflaatio voitaisiin määritellä.

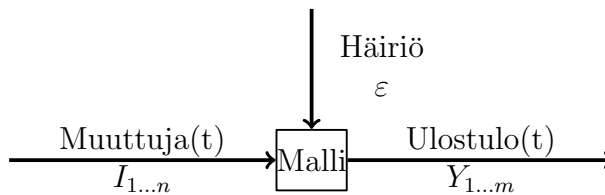
Datan keräyksessä käytetyt määrittelyt voivat vaihdella riippuen kerääjästä. Esimerkkinä kauppaylijäämä kahden valtion välillä voi vaihdella riippuen mittaaajan määrittelystä[11]. Myös ennustusmallien paremmuus riippuu käytetystä mittarista. Näin eri mittareilla mitattuja malleja ei voida verrata keskenään jos ei tunneta mittareiden välistä muunnosta.

Datan esikäsittelyllä tarkoitetaan tässä yhteydessä ennen varsinaista mallintamista datalle suoritettavia toimenpiteitä, joilla yritetään poistaa ylläolevia epätäydellisyyksiä datasta. Datassa olevien mahdollisia virheitä voi olla lukematon määrä ja ne liittyvät kiinteästi käytettävän datan aihealueeseen.

2.4 Mallin tyyppi

M3-ennustuskilpailun tuloksista on todettu ettei tilastollisesti hienostuneempi ja monimutkaisempi ennustusmalli välttämättä tuota tarkempia tuloksia kuin yksinkertaiset menetelmät[?]. Myös menetelmien väliset suhteelliset paremmuusjärjestykset vaihtelevat riippuen mitä tarkkuusmittaria käytetään ja kuinka pitkälle ennustetaan[?]. Yhdistetyillä ennustumalleilla saadaan keskimäärin tarkempia ennusteita kuin yksittäisillä malleilla[?]. Toisaalta esimerkiksi eksponenttiaalimallinnuksella on yksistään pidemmän aikaa saavutettu tarkkojakin tuloksia [12][13][14]. Eksponenttiaalimallia käytettäessä tuotantoketjun mallintamiseen on kumuloituvan virheen vaara, jolloin pieni äkillinen loppukysynnän vaihtelu kumuloituu liioitellun suureksi ennusteeksi tuotantoketjun alkupäässä[15][16][17].

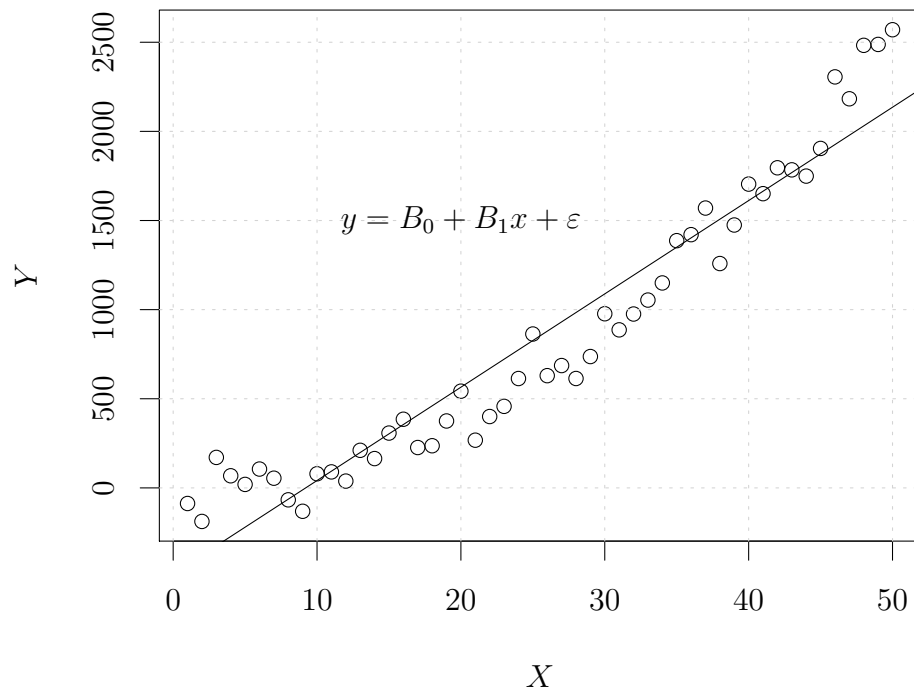
Prosessia mallinnettaessa pyritään löytämään sisääntuloiksi keskenään riippumattomia mutta ulostuloon riippuvia muuttujia [9]. Mallilla pyritään kuvaamaan ulostulon riippuvuus sisääntuloista (Kuva 2.5). Tunnettuja häiriölähteitä voidaan esisuodattaa ennen mallintamista tai mallinnuksen aikana.



Kuva 2.5 Matemaattisen mallin sisääntuloiksi $I_{1...n}$ valitaan muuttujia, jotka ovat keskenään riippumattomia mutta riippuvia ulostulon $Y_{1...m}$ kanssa. Mallilla pyritään kuvaamaan ulostulon riippuvuus sisääntuloina toimivista muuttujista.

Käytettävän mallin valinta riippuu saatavilla olevasta datasta, systeemin luonteen oletuksista, käytettävistä resursseista sekä mallin toimivuudesta kyseiseen tilanteeseen. Tässä luvussa esitellään lineaarinen, epälineaarinen ja syklinen mallinnus. Malleja on lukeamattomia ja usein ne ovatkin monien erityyppisten mallien sekoituksia, joita tarkennetaan mallin eri iteraatiokerroilla.

Lineaarisessa mallinnuksessa muuttujien väliset suhteet ovat lineaarisia. Mikäli systeemiä käytetään jollakin rajatulla toimialueella, voidaan epälineaaristakin systeemiä aproximoida lineaarisilla malleilla (Kuva 2.6) ja saavuttaa riittävä toiminnal-



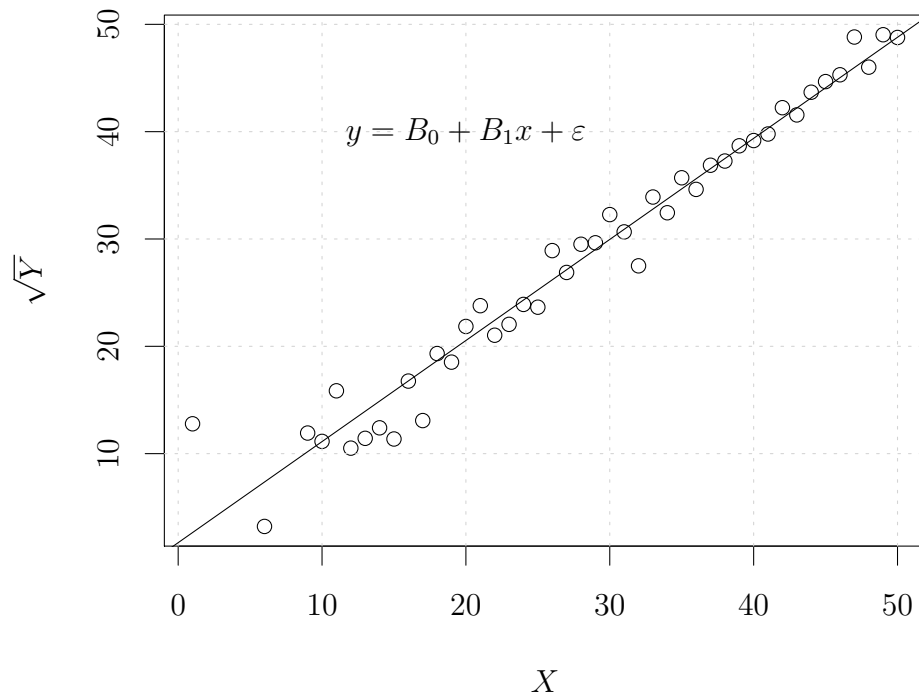
Kuva 2.6 Epälineaarisesti käyttäytyvän systeemin mittauspisteisiin on sovitettu suora virheen pienimmän neliösumman menetelmällä.

linen tarkkuus. Lineaarisen yhtälön yleinen muoto on

$$y = B_0 + B_1x + B_2x + \dots + B_nx + \varepsilon \quad (2.7)$$

Epälineaarinen mallinnus kuvaa sisäänmenojen epälineaarisia suhteita. Oppivat mallit, kuten kohonen-malli[18], voivat muuttaa muuttujien välisiä riippuvuussuhteita ajan mukaan. Tätä suhteiden mukautumista kutsutaan oppimiseksi. Mikäli oppivia malleja käytetään ennustamiseen, täytyy huomioida ettei niiden tarkkaa käyttäytymistä voida koskaan tietää ennakolta mallin toiminnan muuttuessa jatkuvasti. Epälineaarisia systeemejä mallinnettaessa Hyndman suositteli [5] seuraavia keinoja:

1. Voidaanko muuttujien suhde laskea tai muuten päätellä?
2. Saadaanko muuttujat sopimaan suoralle ottamalla niistä muunnoksia?
3. Voidaanko suhteeseen sovittaa jokin malli?



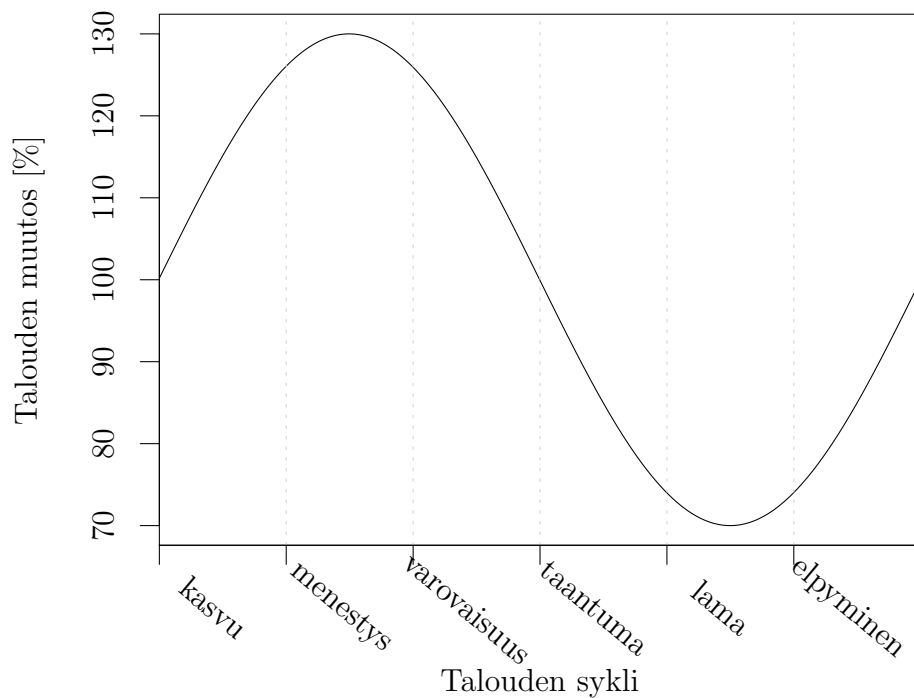
Kuva 2.7 Muuttujalle Y on tehty neliöjuurimuunnos ennen suoran sovittamista virheen pienimmän neliösumman menetelmällä.

Suhteen päättelyyn voidaan käyttää apuna tunnettuja fysikaalisia riippuvuussuhteita. Esimerkkeinä toiselle tai molemmille muuttujille tehtävistä muunnoksista ovat logaritminen- tai neliöjuurimuunnos. Suhdetta voidaan hakea tietokoneavusteisesti iteroimalla gnm- tai nlstools -pakettien[19][20] avulla. Kuvassa 2.7 on otettu neliöjuurimuunnos muuttujasta Y , jonka jälkeen suora sopii mittauspisteisiin paremmin kuin aikaisemmin kuvassa 2.6.

Chu ja Zhang vertailivat lineaarisien ja epälineaaristen mallien toimivuutta vähittäismyynnin ennustamisessa [21]. Vähittäismyynnin ollessa tyypillisesti voimakkaasti kausittaista käyttivät he perinteisiä lineaarisia kausiennustuksen metodeja, kuten aikasarja ja regressio -menetelmiä kiinteillä muuttujilla ja kausikomponentteihin jakoa. Epälineaarisista menetelmistä he soveltivat neuraaliverkkoja, jotka ovat yleistetty epälineaarisilla funktioaproximaattoreilla. Vähittäismyynnin ennustuksessa he saivat parempia tuloksia epälineaarisilla menetelmillä. Paras menetelmä oli aikasarjan analysointi neuraaliverkoilla komponentteihin jaon jälkeen. Lineaariset menetelmät saattoivat tuoda myös hyviä tuloksia, jotka eivät kuitenkaan heidän tutkimuksissaan

toimineet riittävän luotettavasti. Stojanović käytti myyntiä mallintaessaan rakentamaansa lineaarista mallia[22]. Hän otti huomioon myytävän tuoterakenteen muutokset kokonaismyynnin lisäksi. Näin oli mahdollista tarkastella myös tuotteiden keskinäistä käyttäytymistä ja tehdä päätöksiä koskien tuotteiden elinkaaria.

Sykliässä mallinnuksessa talouden oletetaan toistavan samoja vaiheita (Kuva 2.8). Jokaiseen vaiheeseen kuuluvat erityispiirteet taloudessa ja yrityksen toiminnassa, joita voidaan käyttää muuttujina mallia rakennettaessa. Hitaampien talouden syklien ja nopeampien kausivaihteluiden raja ajallisesti vedetään yleensä yhteen vuoteen. Syklejä ennustettaessa keskitytään suuriin kokonaisuuksiin, kuten toimialaan tai tuotteen elinkaareen. Mikäli yritys toimii usealla toimialalla siihen vaikuttavat useat eri syklit. [23]



Kuva 2.8 Taloutta voidaan mallintaa syklien avulla. Jokaiseen syklin vaiheeseen kuuluu omaispiirteitä, joita tutkimalla voidaan tunnistaa nykyinen vaihe. Koellinger ja Thurik jakoivat syklin kuuteen osaan: kasvu, menestys, varovaisuus, taantuma, lama ja elpyminen. Muokattu lähteestä[23].

Koellinger ja Thurik tutkivat 22 eri OECD-maassa talouden syklien vaikutusta yrittäjyyteen vuosien 1972-2007 välillä[24]. Tutkimuksessa todettiin yrittäjyyden kasvun toimivan muutoksen voimana jopa voimakkaalle taloudelliselle kehitykselle.

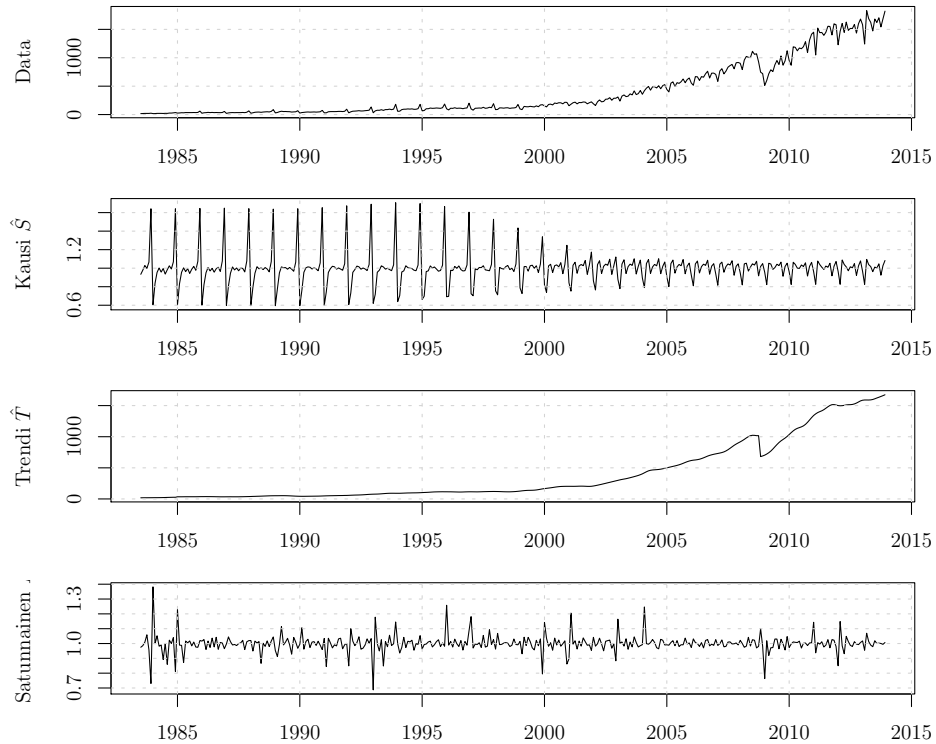
Toisaalta yrittäjäyys lisääntyy vain tilanteissa missä töitä ei ole normaalisti saatavilla. He näkivätkin yrittäjäyden kasvun merkkinä siirtymisestä talouden syklissä lamas-ta elpymisvaiheeseen (Kuva 2.8). Syklisessä mallinnuksessa ajatellaan tapahtumien vaativan oikean ajoituksen tapahtuakseen. Stephanie Schmitt-Grohé ja Martín Uribe1 tutkivat vektoripohjaisen mallinnuksen toimivuutta sykliseen talouden ennus-tukseen ja päätyivät johtopäätökseen että noin puolet talouden vaihteluista johtuvat ennakoitavissa olevista sykleistä [25]. Maximo Camacho ja Jaime Martinez-Martin tutkivat Markov-switching dynamic factor mallin toimivuutta markkinasyklien en-nustamisessa kvartaaleittain [26]. He totesivat mallin olevan käyttötarkoitukseensa sopiva. Michael Halling, Jin Yu ja Josef Zechner tutkivat syklien vaikutusta yk-sittäisiin yrityksiin [27]. He totesivat firmojen toteuman olevan selvästi parempi menestyksen aikana kuin laman. He eivät pystyneet jaottelemaan yrityksiä suhdan-neriippuvaisiin ja -riippumattomiin sillä kaikista oli löydettävissä molempia puolia. Tämän he uskoivat johtuvat yrityksisäisistä rakenteista, joten suhdanteista ei voi suoraan ennustaa yrityksen toteumaa.

Komponentteihin jako tarkoittaa signaalin jakamista osiin, joiden oletetaan käyttäytyvän itsenäisesti. Komponenttien jakoon voidaan käyttää lineaarisia tai epälineaarisia me-netelmiä. Luvussa 2.4.1 on esitelty lineaarinen x11-menetelmä komponenttien muo-dostamiseen.

Mallintajan täytyy tehdä ennakkoratkaisu millaista mallia lähtee ensin kokeilemaan. Tähän vaikuttaa niin saatavilla oleva data kuin käytettävät resurssitkin. Vaikka tut-kimuksia ja sovelluskohteita löytyy paljon ennustamisesta ei niiden tarkkuuksista voi vetää johtopäätöksiä mallien toimivuudesta käytettävälle datalle. Mallien toi-mivuutta käytössä olevalle datalle voidaan arvioida vain sovittamalla malli, simu-loimalla tilanne ja mittaamalla tarkkuus. Mallin todelliset suoritusarvot ilmenevät vasta ajan kuluessa seuraamalla sen käyttäytymistä todellisessa tilanteessa. Yksin-kertaisellakin mallilla voidaan saavuttaa riittävä tarkkuus mikäli mallintajalla on hyvä asiaosaaminen oleellisista syy- ja seuraussuhteista.

2.4.1 X11-komponentteihin jako

X11-komponentteihin jako -mentelmän kehtitti alunperin US Census Bureau and Statistics Canada [5]. Sen avulla voidaan jakaa annettu arvosarja komponentteihin



Kuva 2.9 Kiinan tuonti kuukausittain 1983/7 - 2003/11 jaoteltuina kausi-, trendi- ja satunnaiskomponentteihin x_{11} -menetelmän avulla. Data ja x_{11} -menetelmän implementointi [28].

(Kuva 2.9), jotka voivat olla

$$\hat{Y}_{m_t} = \hat{P}_t + \hat{D}_t + \hat{T}_t + \hat{S}_t + \hat{C}_t + \hat{I}_t \quad (2.8)$$

, missä \hat{Y}_{m_t} =summattu sarja, \hat{P}_t =kuukausivaikutukset, \hat{D}_t =kauppapäivien vaikutukset, \hat{T}_t =trendi, \hat{S}_t =kausittaisuusvaikutukset ja \hat{I}_t =epäsäännöllisyydet [6]. Komponentteihin jako koostuu viidestä pääosasta: kauppapäivien korjauksesta, trendin estimoinnista liukuvalla keskiarvolla, kausittaisten kertoimien vaikutuksen ja soveltamisen valmistelusta, poikkeavien arvojen käsittelystä ja komponenttitaulukkojen ja tilastollisten yhteenvedojen luonnista.[6]

Kauppapäiviä on eri määrä eri maissa ja kuukausissa, joten ensimmäinen vaihe X11-proseduurissa on kauppapäivien määrien korjaus. Tämä tehdään laskemalla kuukausille päivät tasaavat kertoimet. Näin saadaan keskenään vertailukelpoiset kuukaudet. [6]

Trendin estimointi tehdään liukuvalla keskiarvolla. Tämä tapahtuu vertailemaalla trendin \hat{T}_t ja epäsäännöllisyyksien \hat{I}_t suhdetta (Kaava 2.8). Ensimmäisen kerran suhde lasketaan 13-yksikön pituisella liukuvalla keskiarvolla, jonka tulos jaetaan korkeaan, keskimääräiseen tai alhaiseen suhteeseen. Korkeille suhteille käytetään 23 yksikön mittaista liukuvaa keskiarvoa, keskimääräisille suhteille 13 yksikön mittaista ja matalille 9 yksikön mittaista liukuvaa keskiarvoa. Poikkeuksena tästä on kvartaalien laskenta, joka toteutetaan 5 kvartaalin liukuvalla keskiarvolla. [6]

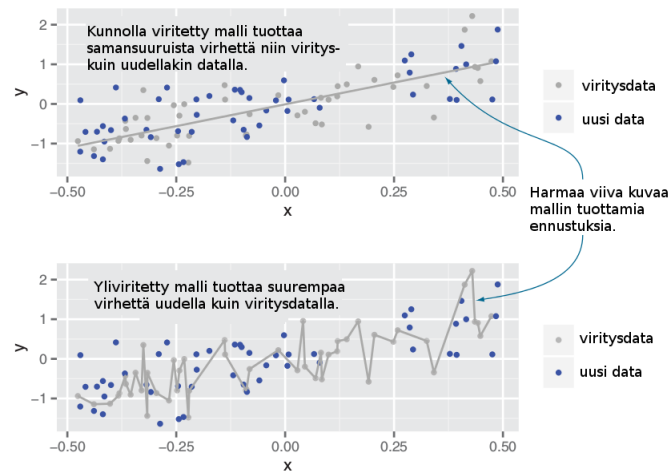
Kolmannessa vaiheessa lasketaan alustavat kausikertoimet, joista lähdetään iteroimalla hakemaan kausikertoimia. Huomioon iteroinnissa otetaan Trendi \hat{T}_t , Kausi \hat{P}_t ja Epäsäännöllisyys \hat{I}_t . [6]

Poikkeavat havainnot (Eng. *outliers*) poistetaan suodattamalla. Tämä tehdään laskemalla esimerkiksi viiden vuoden keskiarvosta tarpeeksi poikkeavat yksiköt. Yleensä tämä tarkoittaa yli $\pm 2.5\sigma$ verran keskiarvosta olevia arvoja. [6]

Viimeisessä vaiheessa luodaan komponenttitaulukot, yhteenvetomittaukset sekä tilastolliset testit varmistuksena komponentteihin jaosta. Adjoin-testissä jokaista kuukautta verrataan edellisen ja seuraavan kuukauden keskiarvoon. Onnustuneessa kausimallinnuksessa suhde ei juuri muutu. Tammikuun testissä verrataan jokaista kuuta tammikuuhun prosentteina. Mikäli näin saadussa sarjassa ei ole suurta vaihtelua, voidaan kausimallinnusta pitää onnistuneena. Tasapainotestissä verrataan 12 kuukauden korjatun sarjan keskiarvoa korjaamattoman vastaavaan keskiarvoon. Mikäli tämä on hyvin lähellä 100% ei sarjaan kohdistu merkittävää kausikorjausta. Testin näyttäessä alle 90% tai yli 110% on mallissa mahdollisesti tapahtunut kausikomponentit ylikorjausta. Korjattuja kuukausia verratessa prosentteina korjaamattomiin voidaan tarkastella korjauksen suuruutta prosentteina. Tällöin voidaan helposti nähdä korjauksen suuruus. [6]

2.5 Ylivirittäminen

Malli on yliviritetty (Kuva 2.10) mikäli on löydettävissä jokin toinen malli, joka suoriutuu huonommin viritysdatalla ja paremmin uudella datalla [29] [30] [31]. Ylivirittämistä ei siten ole aina helppo välittömästi havaita [32]. Tästä syystä eri ennustustapoja vertaillessa pitäisikin tarkastella mallien tuottamien ennustuksien tarkkuuksia (Eng. *prediction interval*) mallien sisäisten tilastollisten tarkkuuksien (Eng. *confidence interval*) sijasta [32]. Mallin ylivirittäminen pyritään välttämään

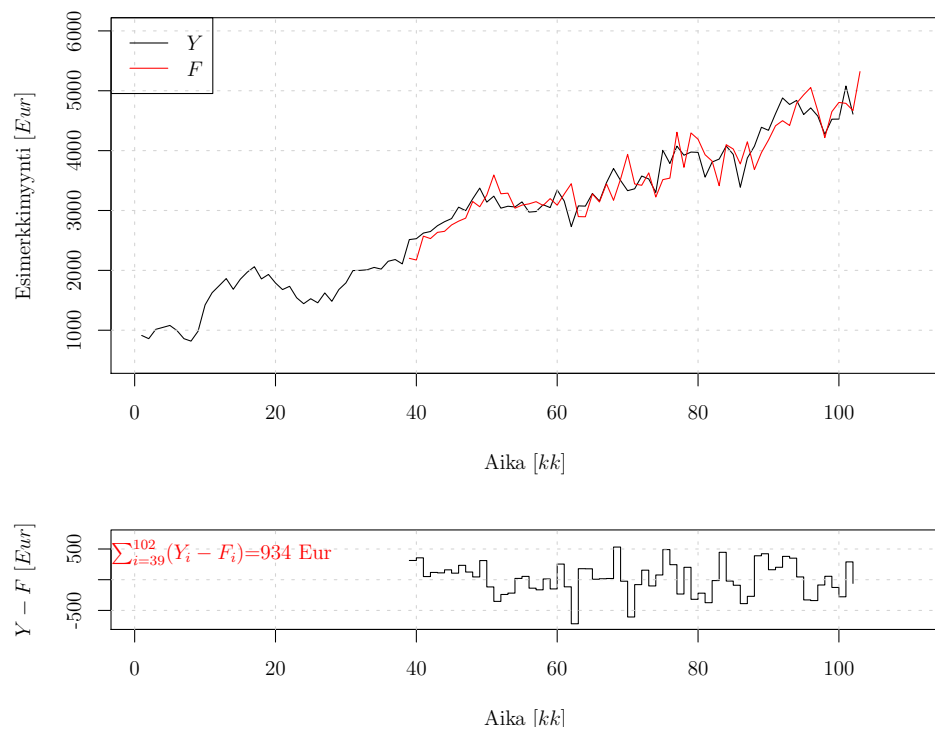


Kuva 2.10 Yliviritetty malli tuottaa tarkempia tuloksia viritykseen käytetyllä datalla kuin uudella. Alemmassa kuvassa malli on yliviritetty. Muokattu lähteestä [30].

käyttämällä eri dataa viritykseen ja mallin tarkkuuden mittaamiseen.

3. TARKKUUDEN MITTAUS

Mallin suoritusarvojen mittaus ja analysointi ovat oleellinen osa mallinnusta, jotta mallin toimivuutta voitaisiin parantaa. Mallin suoritusarvot pyritään mittaamaan niin todenmukaisessa mittausjärjestelyssä kuin käytettävissä olevat resurssit, yleinen turvallisuus ja muut asetukset sen mahdollistavat. Tässä luvussa esitellään mallin



Kuva 3.1 Satunnaiskululla muodostettu esimerkkimyynti Y ja sitä vastaava kuukausiennustesarja F . Jokainen ennuste on laskettu sitä aikaisemmista tiedoista. Alempana on todellisen Y ja ennustetun F sarjan erotus sekä niistä laskettu kumulatiivinen summa.

suoritusarvojen mittaamiseen yleisesti kuuluvat vaiheet: tarvittavan koejärjestelyn rakentaminen, mitattavien suoritusarvojen valinta, mittausajan valinta, virheanalyysi ja tulosten tulkinta. Esimerkkinä toimii satunnaiskululla muodostettu esimerk-

kimyynti, josta on muodostettu kuukausittain virittämällä ja ennustamalla kuukausiennustesarja (Kuva 3.1). Esimerkissä käytettyä x11-ennustemallia (Kappale 2.4.1) käytetään tässä tutkimuksessa myös oikeaan myyntihistoriadataan.

3.1 Koejärjestely

Koejärjestely suoritetaan, joko lopullisessa mallin käyttöympäristössä valvotusti tai sitä mahdollisimman hyvin kuvaavassa tilanteessa. Laivan käyttäytymisj ominaisuuksia on mahdollista testata pienoismallilla tai tietokoneavusteisesti simuloimalla. Lopulliseen käyttöympäristöön malli implementoimalla saavutetaan parhaimmat mittaus tulokset, sillä silloin ei koejärjestelyn epätarkkuuksia tarvitse huomioida erikseen. Tällaista testausta eivät yleensä käytettävissä olevat resurssit, yleinen turvallisuus ja muut asetukset mahdollista ennen huolellista simulointia.

Tšernobylin 26.4.1986 tapahtuneen ydinvoimalaonnettomuuden perimmäinen syy oli virheellisesti rakennettu reaktori, jonka ongelmat realisoituivat turvalaitteiden testaamisesta vastoin määräyksiä käynnissäolevaan reaktoriin[33]. Kokeen suorittaminen todellisessa ympäristössä tuotti tarkan kuvan testattavan turvallisuusjärjestelmän toiminnasta mutta aiheutti samalla laajoja terveyshaittoja sivullisille.

Käytettävissä olevat eri resurssit rajaavat testaamista. Ajallisesti vuosia kestäneen tutkimusaineiston keräämistä ei ole aina mahdollista uusaa vaikka myöhemmin huomattaisiinkin parempi tapa aineiston keruuseen. Käytettävissä oleva osaamis- ja rahoituspääoma vaikuttavat mahdollisten koejärjestelyiden tekemiseen. Tämä voi näkyä verrattavien mallien määränä ennustusjärjestelmiä verrattaessa. Joillakin tieteenaloilla myös etiikka on vahvempana läsnä mahdollisia kokeita suunniteltaessa, kuten oli Dolly lampaan kloonauksesta vuonna 1997 [34].

Ennustusjärjestelmien koeympäristöä rakennettaessa täytyy ottaa huomioon että vain ennustushetkellä olevia tietoja voidaan hyödyntää jokaista ennustetta laatiessa. Ennusteen tarkkaa tarkkuutta ei voida ilmoittaa ennustetta laatiessa vaan se voidaan laskea halutulla mittarilla vasta ennustun ajan jälkeen. Ennustuksien paikkaansa-pitävyyksien toteaminen todellisessa ympäristössä vie niin kauan aikaa kuin kuinka pitkälle tulevaisuuteen ennusteet ovat laadittu. Tästä syystä simuloimalla historiallisia ennustuksia voidaan nopeuttaa mallin tarkkuusanalyysia huomattavasti.

3.2 Arvosarjan osittelu

Mittausfrekvenssi tarkoittaa systeemistä kerättävän datan näytteenottotiheyttä. Liian tiheä frekvenssi tuottaa liikaa tietoa, joka vaikeuttaa datan jatkokäsittelyä ja siten myös analysointia. Liian harva näytteenottotaajuus puolestaan hävittää tietoa systeemin käyttäytymisestä. Esimerkkimyynnissä (Kuva 3.1) mittausfrekvenssi on yksi kuukausi. Maximo Camacho ja Jaime Martinez-Martin saivat kvartaaliennusteita tehdessään hyviä tuloksia datan ollessa kuukausittaista [26], joten mittausfrekvenssin ei tarvitse olla sama kuin ennustusfrekvenssi.

Analysointijaksolla tarkoitetaan aikaa, jolloin mallin suoriutumista mitataan. Suhteellisesti lyhyellä analysointijaksolla tapahtuvien yksittäisten mittaustulosten vaikutus tarkkuuteen on suurempi kuin suhteellisesti pitkällä jaksolla. Analysointijakson pituuden valinta riippuu mallin lopullisesta käyttötarkoituksesta. Jotta ennustetarkkuutta voitaisiin mitata tilastollisin mittarein, tarvitaan useita ennusteita analysoitavan ennustesarjan muodostamista varten. Esimerkissä (Kuva 3.1) analysointijakso on asetettu 39-102 kuukausien väliin, jonka aikana ennusteiden ja tapahtuneen myynnin eroa voidaan verrata eri mittareilla. Viimeisimmän kuukauden ennusteen tarkkuutta ei voida mitata ennen sen realisoitumista. Michael Clements vertaili tekijöihin perustuvaa ennustusta yksikertaiseen AR-ennustusmenetelmään (Eng. *Autoregressive*) [35]. Paremmuus riippui valitusta analysointijaksosta riippumatta olisiko kyseessä reaaliaikainen ennustus vai ei. Tekijöihin perustuva ennustusmenetelmä suoriutui paremmin analysointijakson sisältäessä 1970-luvun finanssikriisin.

Virityspituus on mallin viritykseen käytetyn historiallisen datan pituus. Tämä voidaan rajata alkamaan saatavilla olevan arvosarjan alusta ja loppumaan juuri ennen tarvittavaa ennustusta. Tällöin oletetaan että koko saatavilla oleva arvosarja kuvaa nykyistä ennustettavaa systeemiä. Mikäli saatavilla olevan historian aikana on tapahtunut riittävän suuria muutoksia, kuten yritysfuusioita tai muita markkinamuutoksia, ei kaiken historiatiedon käyttäminen paranna välttämättä ennustetuloksia. Esimerkissä (Kuva 3.1) virityspituus alkaa jokaista ennustetta varten alusta ja loppuu juuri ennen kyseistä ennustetta.

Viritykseen on käytettävissä jokainen kuukausi lisää historiatietoa. Tästä voidaan tehdä oletus ennustetarkkuuksien jatkuvasta paranemisesta historiatiedon lisääntyessä. Robert Rieg tutki tämän hypoteesin paikkaansapitävyyttä kolmen automallin myynnin kuukausiennustetarkkuuksissa kuudessa maassa 15 vuoden aikana [36]. Lopputuloksena hän totesi hypoteesin vääräksi eli tutkimus ei osoittanut ennustetark-

kuuksien parantuneen ja oli siitä yllättynyt vaikka matemaattiset menetelmät olivat parantuneet huomattavasti kyseisenä aikana. Rieg epäili epävarmuuden johtuvan toialan yleisestä epävarmuudesta ja vaihtelusta eikä yrityksen hallinnassa olevista tekijöistä.

Ristiintarkastuksessa (Eng. *Cross-validation*) ennustetaan vuorotellen pientä arvosarjan osaa. Tällä tapaa voidaan saavuttaa paremmin viritetty malli mikäli saatavilla olevaa historiatietoa on vain vähän käytettävänä[37]. Lisäksi mallista saatava tarkkuus kuvastaa toimintaa koko historian aikana [38].

3.3 Suoritusmittari

On tärkeä muistaa ettei ole olemassa parasta ennustusmenetelmää, kuten ei ole olemassa parasta äänentoistojärjestelmääkään. Näiden molempien valintaan vaikuttavat käyttäjän tarpeet ja käytössä olevat resurssit. Tästä syystä pitäisikin muistaa kiinnittää huomiota eri vaihtoehtojen välisiin eroavaisuuksiin, jotta käyttäjä voisi itse tehdä rationaalisen päätöksen sopivimmasta vaihtoehdosta.[39]

Ennustetun sarjan vastaavuus todelliseen arvosarjaan mielletään tarkkuudeksi (Kuva 3.1). Ei ole olemassa vain yhtä sovittua mittastandardia, joilla arvosarjojen eroa voitaisiin mitata. Goojer ja Hyndman tutkivat ennustamistapojen kehittymistä vuosien 1982-1985 Journal of Forecasting:ssa ja vuosien 1985-2005 International Journal of Forecasting:ssa julkaistujen ennustetutkimusten avulla [32]. Osana työtään he luettelivat kyseisissä tutkimuksissa yleisimmin käytetyt numeeriset tarkkuusmittarit (Taulukko 3.1) arvosarjojen vertailuun.

Taulukossa 3.2 on esitetty esimerkkiennusteen tarkkuus yleisimmin käytetyillä mittareilla lisättynä ABB-mittarin keskiarvolla MABB. Jokainen mittari kertoo yhden näkökulman ennusteen tarkkuuteen, joten useamman mittarin avulla on mahdollista saada parempi ymmärrys virheen luonteesta. Mittarilla MABB esimerkin ennustetarkkuus on lähellä täydellistä $MABB = 100\%$ tarkkuutta mutta MAPE-mittarilla virheen suuruus on $MAPE = 5.87\%$. Eri mittareilla mitattuja tarkkuuksia ei voida suoraan verrata toisiinsa.

Taulukko 3.1 Yleisesti käytettyjä mittareita ennustustarkkuuksien vertailuun. F_t on ennuste ja Y_t todennettu tapahtuma. Muokattu lähteestä [32]

MSE	Keskineliövirhe	$\text{ka}(e_t^2)$
RMSE	Keskineliövirheen juuri	$\sqrt{\text{MSE}}$
MAE	Absoluuttinen keskivirhe	$\text{ka}(e_t)$
MdAE	Absoluuttinen mediaanivirhe	$\text{Md}(e_t)$
MAPE	Absoluuttinen keskivirhe prosentteina	$\text{ka}(p_t)$
MdAPE	Absoluuttinen mediaanivirhe prosentteina	$\text{Md}(p_t)$
sMAPE	Symmetrinen abs. keskivirhe prosentteina	$\text{ka}(2 \frac{ Y_t - F_t }{Y_t + F_t})$
sMdAPE	Symmetrinen abs. mediaanivirhe prosentteina	$\text{Md}(2 \frac{ Y_t - F_t }{Y_t + F_t})$
MRAE	Suhteellinen abs. keskivirhe	$\text{ka}(r_t)$
MdRAE	Suhteellinen abs. mediaanivirhe	$\text{Md}(r_t)$
GMRAE	Geometrisen suhteellinen abs. mediaanivirhe	$\text{gMd}(r_t)$
RelMAE	Absoluuttisten keskivirheiden suhde	MAE/MAE_b
RelMSE	Keskineliövirheiden suhde	MSE/MSE_b
RelRMSE	Suhde keskineliövirheiden juurista	$\text{RMSE}/\text{RMSE}_b$
LMR	Logaritmi keskineliövirheiden suhteista	$\log(\text{RelMSE})$
PB	Prosenttia parempi	$100 \text{ ka}(I\{ r_t < 1\})$
PB(MAE)	Prosenttia parempi (MAE)	$100 \text{ ka}(I\{\text{MAE} < \text{MAE}_b\})$
PB(MSE)	Prosenttia parempi (MSE)	$100 \text{ ka}(I\{\text{MSE} < \text{MSE}_b\})$

$$e_t = Y_t - F_t, p_t = 100e_t/Y_t, r_t = e_t/e_t^*,$$

missä e_t^* on vertailuna käytetyn ennustusmenetelmän virhe.

$I\{u\} = 1$, jos u on tosi ja 0 muuten.

Taulukko 3.2 Esimerkkiennusteen (Kuva 3.1) tarkkuuden mittaus taulukon 3.1 mittareiden ja MABB-mittarin avulla analysointiajan yli mitattuna.

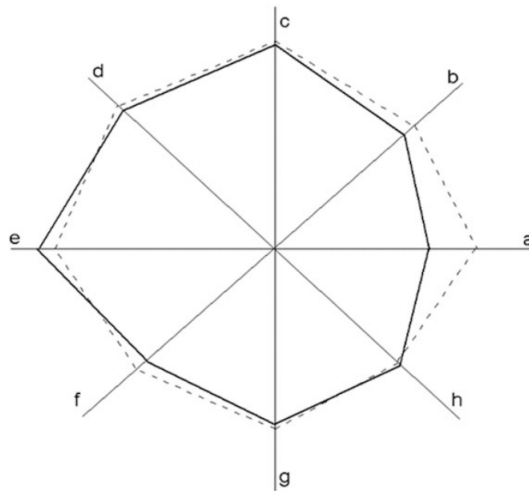
mittari	arvo
MSE	67701.31
RMSE	260.19
MAE	207.80
MdAE	171.04
MAPE	5.87
MdAPE	5.33
sMAPE	0.06
sMdAPE	0.05
MABB	100.90

Everette S. Gardner totesi puolestaan kumulatiivisen virheen summan (Kaava 3.1) olevan käytännöllisin tapa verrata eri ennustusmenetelmien virheitä keskenään [12],

jota käyttivät myös Linder ja Horne tutkiessaan ilmastomuutosta[40].

$$\text{CE}_n = \sum_{t=1}^n Y_t - F_t = \sum_{t=1}^n e_t \quad (3.1)$$

Numeeristen mittareiden lisäksi mallien tarkkuutta voidaan arvioida graafisilla mittareilla. Usealla numeerisella mittarilla mallien vertailun voi toteuttaa käyttäen apuna polar-tyyppistä graafia, jolloin mittarit skaalataan sopimaan samaan kuvioon (Kuva 3.2).



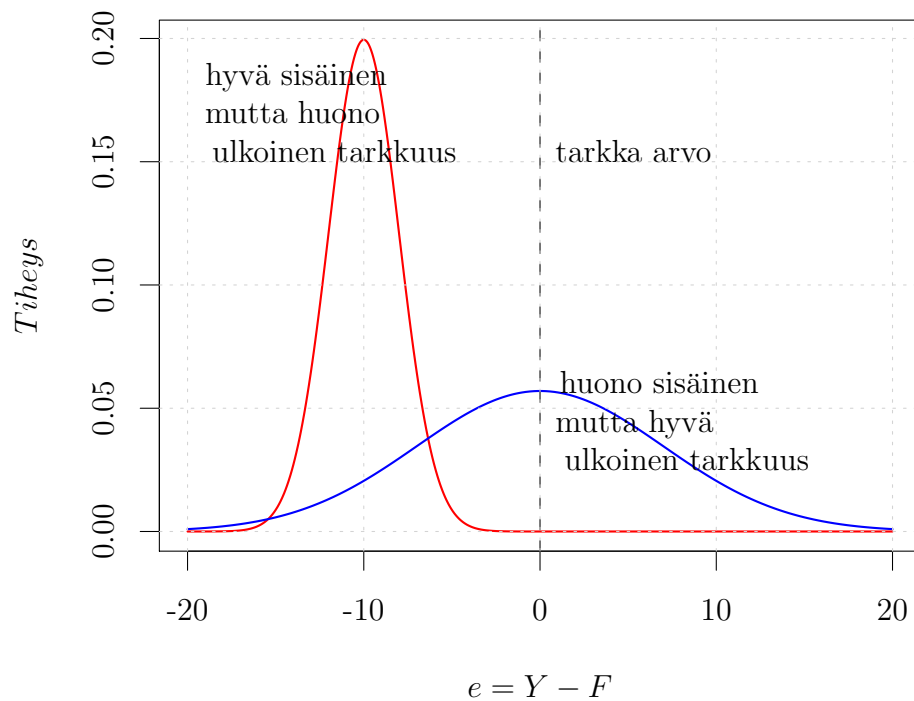
Kuva 3.2 Polar kaaviolla voidaan verrata malleja usean eri mittarin ($a-h$) avulla. [41]

Suhteellista virhettä on helpompi vertailla muihin eri yksikössä oleviin malleihin mutta ne eivät välttämättä toimi hyvin jos suureista toinen tai molemmat ovat nolla tai sen läheisyydessä [41]. Jordi McKenzie tutki epäsymmetrisen **MAPE**-mittarin käytöstä johtuvia ennustevirheitä[42]. Hän totesi tutkimuksessaan että mikäli mittaria käytetään mallin valinnassa valitun mallin ennustuksiin saattaa muodostua systemaattista virhettä. Tämän takia on kehitetty monimutkaisempi symmetristä virhettä mittaava **sMAPE**-mittari. Arvosarjojen vertailuun käytettävän tarkkuusmittarin valinta ja sitä kautta sopivimman ennustusmallin löytäminen ei ole yksiselitteinen tehtävä. Käyttökelpoisimmat ennusteet syntynevät mikäli mittarin valinnassa otetaan huomioon ennusteiden loppukäyttäjien tarpeet.

3.4 Virheanalyysi

A.J. Jaffe ja Herbert F. Spirer jakavat data-analyysissä tapahtuvat virheet viiteen osaan: puute asiaosaamisessa, lähtödatan oikeellisuus (Katso kappale 2.3), tutkimuksen suunnittelu ja esitys, tilastolliset menetelmät ja tahallinen tietojen peittäminen[11]. Yleisien virhepaikkojen lisäksi tässä luvussa esitellään ulkoinen ja sisäinen virhe sekä arvo- ja aikavirhe.

Tarkkuusmittaukset voidaan jaotella sisäisiin ja ulkoisiin tarkkuuksiin [43] (Kuva 3.3). Ulkoinen tarkkuus pyritään mittaamaan mallin ulkopuolelta objektiivisesti, kuten simuloimalla ennustusmallien ennustustarkkuus. Tutkimuksien tuloksia ver-

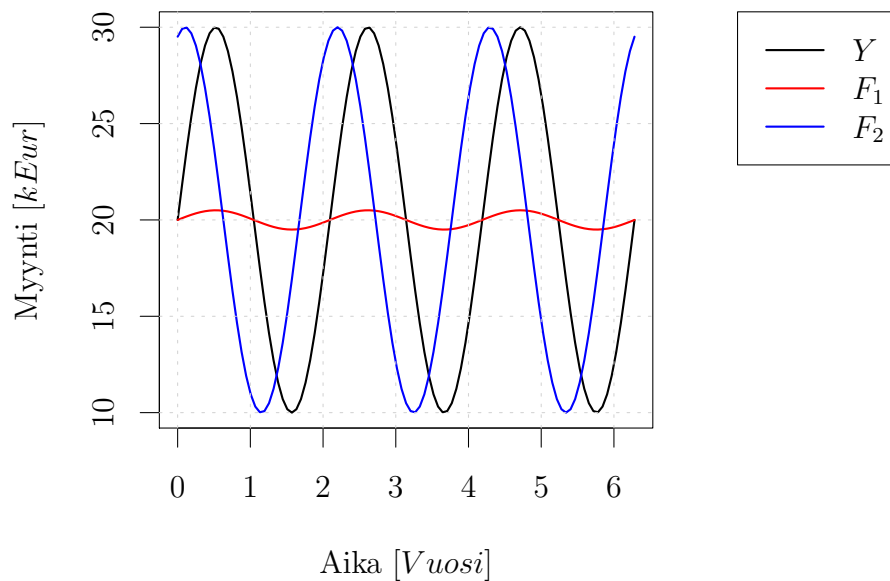


Kuva 3.3 Virhe voi olla sisäistä (sininen) tai ulkoista (punainen). Muokattu lähteestä [43]

ratessa täytyy mittaustapaan kiinnittää huomiota sillä molempia tarkkuuksia voidaan mitata samoilla mittareilla. Kuvassa 3.3 on mitattu kahta mallia, ensimmäistä sisäisesti ja jälkimmäistä ulkoisesti. Ensimmäinen on sisäisesti toista tarkempi mutta jälkimmäinen ulkoisesti ensimmäistä tarkempi. Ensimmäisen mallin virhe on systemaattinen 10 yksikön suuruinen, joka voidaan mittauksen jälkeen korjata. Yleisesti

ottaenkin tunnetut virhelähteet on tuntemattomia helpompi ottaa huomioon.

Virheet voidaan jakaa myös arvo- ja aikavirheisiin. Oletetaan kuvan 3.4 mukainen tilanne, jossa tuotteen todellinen myynti Y on merkitty mustalla ja mallien F_1 ja F_2 tuottamat ennusteet punaisella ja sinisellä. Malli F_1 pystyy ennustamaan tarkasti



Kuva 3.4 Todellinen myynti Y ja mallien F_1 ja F_2 ennustukset. Mallin F_1 ennustuksessa on euromääräinen virhe ja mallissa F_2 ajoitusvirhe.

todellisen myynnin huipun ajankohdan mutta ei tarkkaa myyntimäärää. Malli F_2 kuvaa puolestaan tarkasti myyntimäärää mutta ajoitus on väärä. Taulukkoon 3.3 on laskettu molemmista ennustuksista yleisimmät numeraaliset mittarit (Taulukko 3.1) ja MABB (Kaava 4.2), joiden perusteella tarkkuutta verrataan.

Taulukon 3.3 mittarien tuloksista nähdään jokaisen mittarin puoltavan mallia F_1 . Tämä johtuu siitä että jokainen mittari mittaa arvo- eikä aikavirhettä. On mahdollista löytää tilanteita, kuten riittävän aikavälin budjetin laatiminen, jossa mallin F_2 ominaisuudet voisivat olla eduksi. Mittarien valinnassa pitää huomioida käyttötarkoitus ja olla samalla tietoinen mitä heikkouksia laskentatapoihin mahdollisesti liittyy. Oikeissa ennustuksissa aikavirhe ei välttämättä ole yhtä selvästi esillä.

Taulukko 3.3 Ennustuksien F_1 ja F_2 (Kuva 3.4) tarkkuuden vertailu yleisesti käytössä olevien numeeristen mittarien (Taulukko 3.1) sekä MABB (Kaava 4.3) avulla. Kaikki mittarit puoltavat F_1 mallin tuottamia ennustuksia.

mittari	F_1	F_2
MSE	44.68	69.31
RMSE	6.68	8.33
MAE	5.99	7.50
MdAE	6.50	8.57
MAPE	36.19	42.63
MdAPE	28.94	41.10
sMAPE	0.31	0.40
sMdAPE	0.32	0.43
MABB	99.41	110.26

Tilastollisia menetelmiä voidaan käyttää väärin lukemattomin tavoin. Pelkästään lukujoukosta ilmoitettava tyypillinen luku voidaan laskea keskiarvolla, mediaanilla, moodilla, geometrisella keskiarvolla tai harmonisella keskiarvolla riippuen tutkijan valinnoista. Näin julkistettavia ja erikseen esitettäviä lukuarvoja voidaan valita sopivaksi riippuen käyttötarkoituksesta. Tilastollisten menetelmien hyödyntämisessä voi tapahtua myös inhimillisiä virheitä, joita ei ole välttämättä helppo huomata. Tilastollinen korrelaatio voi esimerkiksi sekoittua kausaalisuuteen tutkimuksien tiivistelmissä. Virheiden tarkistuksen vaikeudesta johtuen ei voida poissulkea mahdollisuutta niiden tekemistä tarkoituksella tuloksien vääristämiseksi tutkijan kannalta suotuisaan suuntaan.

Asiaosaamisen puute voi näkyä virheellisten tilastollisten laskujen lisäksi myös hypoteesin muodostuksessa, tiedon etsinnässä ja lopputuloksien tarkastamisessa. Tutkimuksen suunnittelussa pitää pyrkiä kumoamaan tunnettujen ja tuntemattomien virhelähteiden vaikutukset satunnaisotannalla. Kuvaajien akselien skaalauksilla voidaan vaikuttaa tietoisesti tai tiedostamatta tuloksista välittyvään mielikuvaan. Lisäksi tuloksien merkittävyyydestä tehtävät tulkinnat saattavat vaihdella tulkitsijoiden välillä, joten pelkkien tulkintojen osien esittäminen asiayhteydestä erillään ei välttämättä kuvaa enää tutkimuksen tuloksia. Näitä virheitä vältetään akateemisia tutkimuksia tehdessä valitsemalla vain ensimmäisen käden tutkimuksia ja tutkimalla kriittisesti itse tutkimusta lähtöoletuksineen pelkkien loppupäätelmien sijasta.

4. SIMULOINTIJÄRJESTELMÄ

Ennustusjärjestelmän tarkoituksena on tuottaa automaattinen myyntiennuste perustuen historialliseen tietoon kaikelle, kanavittaisille tai maakohtaisille myynneille ennusteaajan ollessa seuraava kuukausi, kvartaali tai loppuvuosi. Osaan vaadittavista ennustuksista on annettu tarkkuusvaatimus (Taulukko 4.1), joita paremmin ennustettaessa voidaan sanoa järjestelmän täyttäneen ennako-odotukset.

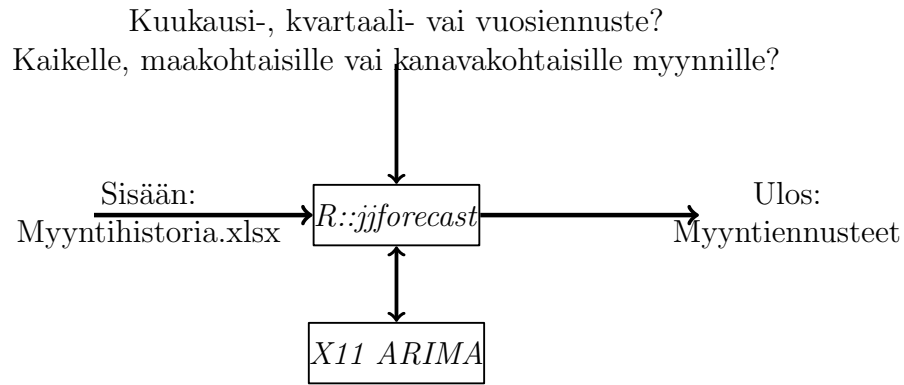
Taulukko 4.1 Ennustusjärjestelmälle asetetut tarkkuusvaatimukset ABB_t -mittarilla (Kaava 4.1). Sulkeissa olevat tarkkuudet ovat lisätty analyysia varten eivätkä siten ole virallisia vaatimuksia.

Data			
Kaikki	Kanavat	Maat	Ennusteaika
80%	70%	50-60%	Kuukausi
90%	[80%]	[60-70%]	Kvartaali
-	-	-	Loppuvuosi

Myyntihistoria oli tallennettu 46.6 MB kokoiseksi **.xlsx*-muotoiseksi tiedostoksi, johon sisältyi jokaisen yksittäisen myyntitapahtuman tiedot, kuten esimerkiksi tuotteen, ostajan ja myyjän tietojen lisäksi myös kyseisen myynnin euromääräinen arvo ja myyntipäivä. Tiedoista selvisi myös minkä yhtiön sisäisen kanavan kautta myynti oli tapahtunut ja mihin maahan. Käytetyn historiatiedon ensimmäinen myyntitapahtuma oli 1.2.2008 ja viimeinen 2.6.2016.

4.1 Tekninen toteutus

Annetuista tiedoista on mahdollista muodostaa erilaisia matemaattisia malleja (Kapale 2.4), joiden avulla voidaan tuottaa ennusteita. Jotta tilanteeseen soveltuvien mallien voitaisiin valita, täytyy niiden tuottamien ennustustarkkuudet mitata ja verrata. Tämä tutkimus kuitenkin keskittyy tarkkuuden mittaamiseen, joten useiden mallien viritys ja vertailu rajautui ulkopuolelle ja ovat mahdollinen jatkotutkimusaihe.



Kuva 4.1 Ennustusjärjestelmän toiminnallinen rakenne. Sisääntulona toimii annettu *.xlsx -tiedosto, josta lasketaan valitut ennusteet näytölle. Taustalla ennustuskoneena käytetään X11-Arima -ennustusohjelmaa. Käyttäjä valitsee halutun ennustusjakson sekä kohteen. jforecast toimii tulkkina käyttäjän ja ennustusohjelman välillä hoitaen datankäsittelyn.

Malliksi valittiin erillinen ennustusohjelma, joka hyödyntää Arima-X11 -menetelmää (Kappale 2.4.1). Ohjelman automaattinen mallinnusproseduuri sopii erilaisten arvosarjojen ennustamiseen. Ennen kyseiseen proseduriin päätymistä myös uudempaa Arima-X13-seats -ennustusproseduuria kokeiltiin, joka kuitenkin pystyi tuottamaan käytetyllä datalla harvemmasta maasta ennusteita kuin x11-ennustusmenetelmä ja jätettiin siksi pois. Tarkkuusvertailuja näiden kahden järjestelmän välillä ei ole vertailtu. Ohjelman ylläpidosta vastaa Amerikan Yhdysvaltojen tilastokeskus United States Census Bureau [44].

Ennustuskone ja simulointijärjestelmä toteutettiin R-kielellä, joka on tilastolaskentaan käytetty laskentakieli [45]. Se muodostuu hierarkisesti paketeista, joten laskentatapojen kehittyessä ne päätyvät nopeasti kaikkien käyttäjien saataville. Asentamalla *Seasonal*-paketti[28] on mahdollista käyttää valittua x11-ennustusmenetelmää. R-kieltä voi käyttää ilmaiseksi kaikilla käyttöjärjestelmillä myös ilman graafista käyttöliittymää.

Rakennettu *jforecast*-paketti toimii tulkkina käyttäjän pyyntöjen, annetun historiatiedon ja ennustusohjelman välillä. Sen toiminnallinen rakenne on kuvan 4.1 mukainen, jossa sisääntulona toimii annettu *.xlsx -tiedosto ja ulostulona näytölle tulostuvat ennusteet. Käyttäjä valitsee haluaako seuraavan kuukauden, kvartaalin vai loppuvuoden ennusteen kaikelle, maakohtaiselle vai kanavakohtaiselle myynnille.

4.2 Ennustusten muodostaminen

Ennustusten tekninen toteutus ohjelmoitiin *jfforecast*-pakettiin. Siihen kuuluu syötetyn datan esikäsittely, suodatus, ennustettavien sarjojen muodostaminen, taustalla käytettävän Arima-x11-ennustusohjelman käyttö sekä sen tuottaminen tuloksien suodattaminen käyttäjää tai myöhempää analysointia varten.

4.2.1 Ennustettavat sarjat

Annettu **.xlsx*-muotoinen tiedosto muutetaan R-kielen ymmärtämään muotoon, jonka jälkeen siitä poistetaan muut kuin mallintamiseen käytettävät tiedot. Näin pienennetään ohjelman vaatimia resurssivaatimuksia. Satunnaiset yli määritellyn arvorajan ylittävät kaupat suodatetaan pois niiden vaikean ennustamisen ja suuren vaikeavuuden vuoksi. Ennustettava myynti koostuu siten niin kutsutusta päivittäisestä perusmyynnistä.

Yksittäisistä kaupoista summataan ennustettavia arvosarjoja valittujen parametrien mukaisesti. Esimerkiksi maiden kuukausiennusteita pyydetessä summataan kaupat maittain ja kuukausittain, jolloin saadaan yhtä montaa ennustettavaa arvosarjaa kuin maita on. Käytetyssä historiatiedossa on kauppojen peruutuksia sekä myyntimäärien muutoksia kirjattu negatiivisina myynteinä. Nämä negatiiviset myynnit eivät välttämättä ole samana kuukautena kuin niihin liittyvä alkuperäinen myynti, joten joissain tilanteissa koko kuukauden myyntisumma voi korjauksen vuoksi kääntyä negatiiviseksi. Maiden myynneistä summatut arvosarjat ovat keskenään eripituisia ja voivat sisältää tyhjiä tai negatiivisia kuukausia.

Arvosarjan sisältäessä negatiivisia kuukausia myös siitä muodostettava ennuste saattaa muodostua negatiiviseksi. Tämän vuoksi jokin yksiselitteinen tapa muodostaa todellinen reaaliymynti negatiivisista myynneistä mallinnusta varten voisi lisätä ennustetarkkuutta joissain tilanteissa. Nämä korjaukset ovat suuressa osassa maita suhteellisen pieni osa myynnistä, jolloin myös niiden vaikutukset ennusteisiin oletetaan pieniksi.

Käytettyyn Arima-X11-ennustusohjelmaan ei voida syöttää vajaita kuukausia tai kvartaaleita, joten arvosarjat eivät voi sisältää vajaiden kuukausien tai kvartaalien myyntisummaa. Kuluvan kuun tai kvartaalin myyntisumma poistetaan, sillä vajaina sitä ei voida käyttää osana mallinnukseen käytettävää arvosarjaa. Olisi mah-

dollista myös ennustaa nykyinen kuukausi tai kvartaali ensin loppuun, jonka ennustusta käytettäisiin virittämisen osana. Tämä monimutkaistaisi ja hidastaisi ennustusten tuottamista ja tarkkuuksien analysoimista huomattavasti. Historiatiedon vähimmäisvaatimuksen ollessa kolme vuotta eli 36 kuukautta oletetaan ettei viimeisen kuukauden myyntitiedolla ole suuri vaikutus mallin tuottamiin ennusteisiin. Kvartaaliennustuksia tehdessä viimeisen kvartaalin osuus kolmen vuoden kvartaaleista on suurempi.

4.2.2 Virityspituus

Viritysos on myyntihistorian osa, jota käytetään ennustuksessa käytettävän mallin muodostamiseen. Virityspituus ilmoittaa viritysoson pituuden, joka ilmoitetaan usein samana yksikkönä kuin ennustettava ajanjaksokin. Käytetyllä ennustusohjelmalla on vähintään kolmen vuoden yhtäjaksoisen historiatiedon vähimmäisvaatimus ennen kuin se pystyy tuottamaan ennusteita.

Virittämiseen käytettävän datan pitäisi kuvata mallinnettavaa systeemiä mahdollisen tarkasti. Saatavilla olevan historiatiedon alkupään markkinoiden toiminta saattaa olla hyvinkin erilaista kuin loppupään, jolloin valittavalla viritysosalla voi olla suuri vaikutus saataviin ennusteisiin.

jforecast-paketti käyttää oletuksena kaikkea paitsi viimeistä vajaata kuluva kuukautta tai kvartaalia mallinnukseen. Viritysoson loppupää päättyy ennustushetkellä viimeiseen täyteen kuukauteen tai kvartaaliin, jota ennusteita simuloitaessa liikutetaan analysointijakson yli. Viritysoson alkupää määrätään virityspituus-parametrillä loppupäästä täysinä ajanjaksoina. Virityspituuden vaikutusta saataviin ennustetarkkuuksiin tutkitaan käymällä läpi kaikki mahdolliset virityspituudet.

4.2.3 Ennustus ja tulostenkäsittely

Todelliseksi ennustuspituuksi muodostuu kaksi kuukautta sillä kuluva vajaata kuukautta ei voida käyttää mallinnuksessa. Käytetty ennustusohjelma ilmoittaa pyydetessä ennusteen todennäköisyyden keskiarvon lisäksi sisäisen mallin tarkkuuden (Eng. *Confidence Interval*), jota ei käytetä tarkkuuden analysoinnissa. Erillisessä tarkkuuden simuloinnissa käytetään ohjelman antamaa ennusteen keskiarvoa. Ohjelman ennuste korjataan nolllaksi mikäli se on negatiivinen myyntilukujen korjauksien vuoksi. Ennustetarkkuuden simuloinnissa Arima-x11-ohjelmaan syötetään

kaikki mahdolliset ennustukset kaikilla mahdollisilla virityspituuksilla ja kerätään tulokset myöhempää analysointia varten talteen. Normaalissa ennustustilanteessa pyydettyt ennusteet tulostuvat näytölle jatkokäsittelyä varten.

4.2.4 Loppuvuoden ennusteen toteutus

Kuluvan vuoden kokonaismyynnin ennustaminen on jaettu tunnettuun ja ennustettuun osaan. Tunnettu osa voidaan laskea annetun datan kokonaisista kuukausista ennen ennustushetkeä. Ennustettu osa ennustetaan kuukausiennustamalla asetamalla ennustehorisontti vastaamaan loppuvuoden pituutta sisällyttäen nykyisen kuluvan kuukauden.

Lopullinen vuoden kokonaismyynnin ennuste muodostuu ennustettujen jakaumien keskiarvojen ja tunnetun osan summasta. Tämän tavan ennustetarkkuutta ei ole analysoitu suuren epätarkkuuden vuoksi. Ennustetarkkuus paranee loppuvuotta kohhti ennustettavan osan pienentyessä ja tunnetun osa suureutuessa.

4.3 Tarkkuusmittarien valinta

Tarkkuusmittarin valinta vaikuttaa oleellisesti vaikutelmaan ennustejärjestelmän hyvyydestä ja toimivuudesta. Simuloidun tarkkuusmittauksen tarkoitus on mitata kuinka tarkkoja kuukausi- ja kvartaaliennusteet ovat ja paranevatko ennustetarkkuudet virityspituutta lisätessä. Virityspituuden vaikutusta ennusteiden tarkkuuteen mitataan vaihtelemalla virityspituutta arvosarjoittain. Numeerisiksi tarkkuuden mittareiksi valitaan e_t , CE_n , ABB_t ja $MABB_n$ mittarit (Kaavat 4.1- 4.4), paremman käsityksen muodostamiseksi vain yhteen mittariin verrattuna.

Yrityksen ennustuksien tarkkuusvaatimukset ovat ilmoitettu mittarilla ABB_t , jonka avulla voidaan verrata yksittäistä ennustusta F_t vastaavaan todelliseen myyntiin Y_t .

$$ABB_t = \frac{\text{Todellinen myynti}}{\text{Ennustettu myynti}} \cdot 100 = \frac{100Y_t}{F_t} \quad (4.1)$$

Jotta voitaisiin tarkastella ennustustavan toistuvaa tarkkuutta, täytyy mittarina käyttää jotakin arvosarjoja vertaavaa mittaria. Tähän käytetään ABB_t mittarin keskiarvoa $MABB_n$, missä n on verrattavien arvosarjojen pituus. Mikäli ennustevirheet ovat symmetrisiä mittarilla ABB_t ja/tai nolla niin mittari $MABB_n$ näyttää täydellistä

100% tarkkuutta. Mittarin käyttö on perusteltua sillä sen perusversio, ABB_t , on yrityksessä sisäisesti tuttu ja ennustetarkkuusvaatimukset ovat annettu sen avulla.

$$MABB_n = \frac{1}{n} \sum_{t=1}^n ABB_t = \frac{100}{n} \sum_{t=1}^n \frac{Y_t}{F_t} \quad (4.2)$$

Akateemisen vertailtavuuden vuoksi valitaan toiseksi mittariksi $MAPE_n$, joka on yksi yleisimmin käytettyistä mittareista (Taulukko 3.1). Keskimääräinen absoluuttinen virhe on nolla mikäli ennusteet ovat täydellisiä. Tämä mittari ei erottele $MABB_n$ tapaan onko toteutunut ollut ennustetta suurempi vai pienempi.

$$MAPE_n = \frac{100}{n} \sum_{t=1}^n \frac{|Y_t - F_t|}{Y_t} \quad (4.3)$$

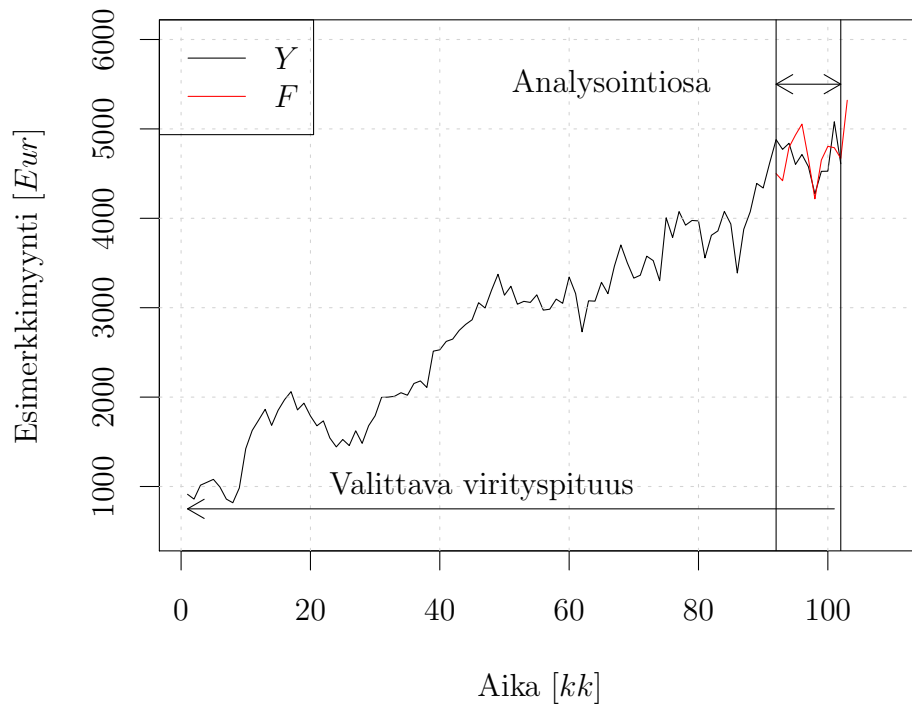
Ennustetarkkuuksia tarkastellessa ilmeni suurimpien prosentuaalisten virheiden tulevan maihin, joissa on suhteellisesti pieni ja vaihteleva myyntimäärä. Tästä syystä kolmanneksi tarkkuusmittariksi on lisätty euromääräinen ennustevirhe e_t tarkastelujakson ajalta ja sen kumulatiivinen summa CE_n tarkkailujakson ajalta. Näin voidaan arvioida kuinka merkittävä kyseinen ennustettavuustarkkuus on euromääräisesti ja kumoavatko virheet euromääräisesti toisensa.

$$e_t = Y_t - F_t \quad (4.4)$$

$$CE_n = \sum_{t=1}^n e_t \quad (4.5)$$

4.4 Tarkkuuksien simulointi ja esitys

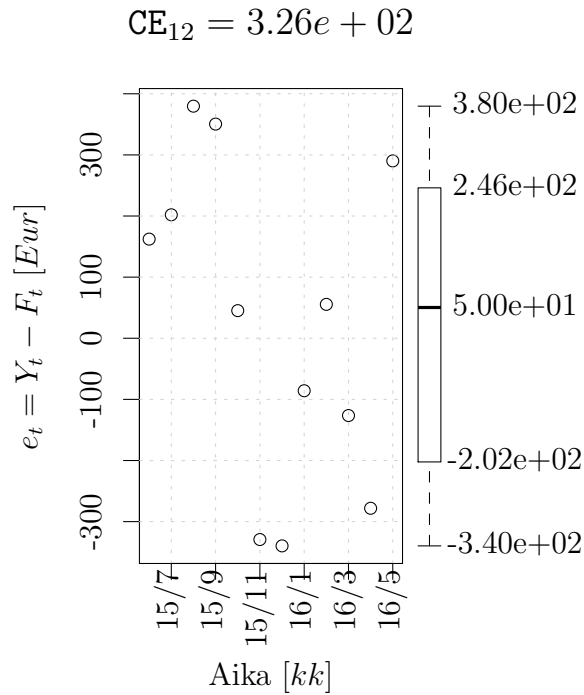
Simuloinnin analysijaksoiksi valitaan 8 kvartaalia kvartaaliennusteille ja 12 kuukautta kuukausiennusteille laskettuna viimeisestä kokonaisesta yksiköstä (Kuva 4.2). Virityspituutta muutetaan mitatessa sen vaikutusta tarkkuuteen. Muulloin se alkaa aina datan alusta ja loppuu viimeiseen täyteen yksikköön ennen ennustetta.



Kuva 4.2 *Analysointiosan ennustuksien tarkkuuden muuttumista virityspituuden funktiona voidaan tarkastella laskemalla ennustukset eri virityspituuksilla. Analysointiosan pituus on kuukausiennustuksille 12 kuukautta ja kvartaaliennustuksille 8 kvartaalia.*

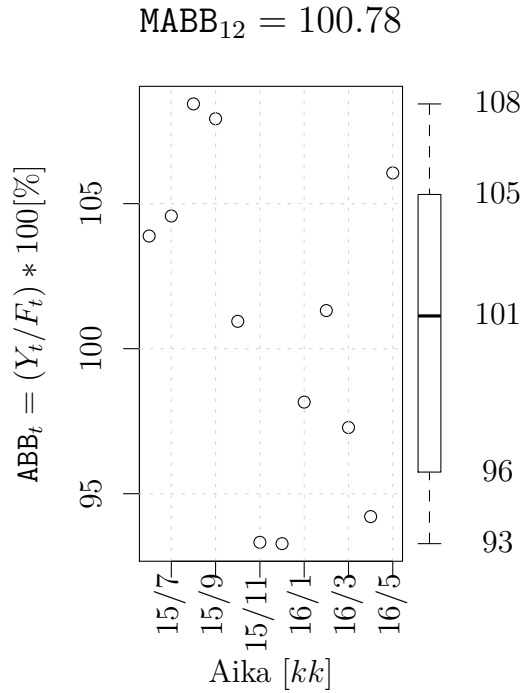
Tarkkuusmittarien, virityspituuden ja analysointijakson valintojen jälkeen voidaan saatavista tarkkuuksista muodostaa graafinen esitystapa. Valitaan kolme eri tarkkuusgraafia (Kuvat 4.3- 4.5) jokaista ennustettavaa sarjaa kohden. Näin voidaan tarkastella käytetyn ennustusmallin vahvuuksia ja heikkouksia erityyppisissä sarjoissa.

Euromääräisistä ennustusvirheistä $e_t = Y_t - F_t$ näkee analysointijakson aikaiset yksittäiset ennustevirheet (Kuva 4.3). Kuvan yläpuolelle on laskettu kumulatiivinen ennustevirhe CE_n , josta nähdään kuinka symmetrisesti virheet asettuvat nollan molemmille puolille. Tästä tiedosta saattaa olla hyötyä esimerkiksi budjetoinnissa. Kuvan oikealla on box-plot, joka antaa toisen näkökulman ennustustarkkuuden jakaumaan euromääräisesti. Box-plotin keskimmäinen arvo on mediaani.



Kuva 4.3 Ennustetarkkuus e_t - ja CE_n -mittareiden sekä box-plotin avulla mitattuna analysointijakson aikana. Ennustusvirhe on mitattu kuvan 4.2 tilanteesta.

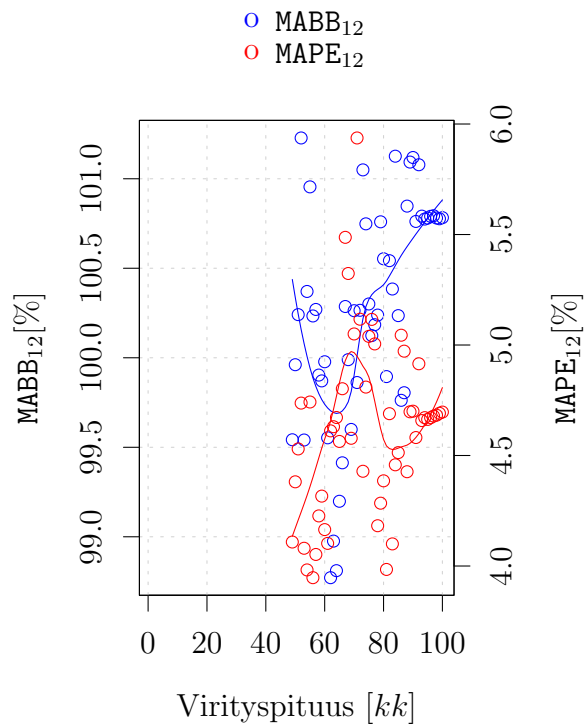
Yhtiön sisäisen tarkkuusmittarin ABB_t avulla voidaan tarkastella yksittäisien ennustusvirheiden suuruutta tarkastelujakson aikana, joista on laskettu keskiarvo $MABB$ kuvan päälle. Box-plot oikeassa reunassa kuvaa virhejakaumaa ABB_t asteikolla. Tämän kuvan perusteella voidaan tarkastella virheiden keskinäistä suuruutta, niiden keski- ja mediaaniarvoja ja vertailla niitä annettuihin tarkkuusvaatimuksiin (Taulukko 4.1).



Kuva 4.4 Ennustetarkkuus ABB_t -ja $MABB$ -mittareiden sekä box-plotin avulla mitattuna analysointijakson aikana. Ennustusvirhe on mitattu kuvan 4.2 tilanteesta.

Analysointijakson tarkkuuden muuttumista virityspituuden funktiona tarkastellaan $MABB_n$ ja $MAPE_n$ mittareiden avulla (Kuva 4.5). Pisteisiin on sovitettu trendi, jota voi käyttää apuna tarkkuuden suunnan tulkitsemisessa. Tämän kuvan perustella voidaan analysoida kahden mittarin avulla miten tarkkuus muuttuu analysointijakson aikana, jos käytetään eri pituuksia historiaa viritykseen. $MABB_n$ -mittarin täydellinen tarkkuus on 100% ja $MAPE_n$ -mittarin 0%. Simuloiduissa ennustetarkkuuksissa voi joko olla selvä trendi tarkkuuden muuttumisesta tai sitten ei. Kuvan 4.5 esimerkissä ei

ole selviä trendejä kummallakaan mittarilla mitattuna, joka on ymmärrettävää ennustettavan sarjan ollessa päivittäisestä satunnaiskulkusarjasta kuukausittain summattu sarja. Kolmen vuoden jälkeen lisätyllä historiatiedolla mallin tuottamat ennustetarkkuudet eivät juuri muutu. Huomattavaa on että mittarien trendit eivät välttämättä kulje samaan suuntaan ja toisen tarkkuuden parantuessa toinen huonontuu.



Kuva 4.5 Ennustetarkkuuden muutos virityspituuden funktiona mitattuna $MABB_{12}$ - ja $MAPE_{12}$ -mittareilla analysointijakson aikana. Virityspituus on laskettu ennustettavaa kuukautta edeltävästä kuukaudesta taaksepäin. Ennustusvirhe on mitattu kuvan 4.2 tilanteesta.

Käytettävällä ennustuskoneistolla on vähintään kolmen vuoden vähimmäisvaatimus historiatiedolle, joten ennustuksia saadaan aikaisintaan 36 kuukauden virityspituudella. Todellisista myyntiluvuista ennusteita simuloitaessa tämä raja vaikuttaa usein olevan 50 kuukauden lähettyvillä, joka näkyy myös esimerkkitalanteessa kuvassa 4.5.

5. ANALYYSI

Kaikkien pyydettyjen kuukausi ja kvartaaliennusteiden tarkkuudet ovat simuloitu ja mitattu esimerkin (Liite ??) tavoin. Loppuvuoden ennusteiden tarkkuutta ei ole tutkittu suuren epävarmuuden vuoksi. Luvussa analysoidaan simulointien tulokset.

5.1 Kuukausiennusteet

Kaiken myynin kuukausiennustuksien tarkkuus oli 2015/6 - 2016/5 tarkasteluvälillä $mean_{12}(abb)$ -mittarilla 99.24% ja $mape_{12}$ -mittarilla 6.64% (Taulukko 5.1). Yksittäiset kuukausiennustetarkkuudet vaihtelivat mittarilla abb välillä 82 – 108%. Ennustus-tarkkuus ei muutu selvästi virityspituuden funktiona mittarilla $mean_{12}(abb)$. Virityspituuden lisääminen tuotti tasaisesti hieman paremman ennustustuloksen mittarilla $mape_{12}$ vertailujakson aikana. Yrityksen tarkkuusmittarin abb kahdentois-ta kuukauden keskiarvo $mean_{12}(abb)$ on hyvin lähellä täydellistä 100% tarkkuutta. Luku kertoo ennustusvirheiden kumoavan toisensa hyvin vertailujaksolla kyseisellä mittarilla. Ennakkovaatimuksena kaiken myynnin kuukausiennustustarkkuudelle oli abb -mittarilla 80%, joka toteutui niin keskiarvollisesti kuin kaikissa yksittäistapauksissakin. Koska riippuen mittarista tarkkuus joko hieman parani tai ei juuri muuttunut virityspituutta lisättäessä, voidaan olettaa muutaman seuraavan kuukauden ennusteen tarkkuuden olevan samaa luokkaa tai jopa hieman edellisiä tarkempi. Kaiken saatavilla olevan historian käyttö on perusteltua molemmilla mittareilla, jotta saataisiin mahdollisimman tarkka ennuste.

Kuukausiennustuksien tarkkuudet neljälle eri kavavalle olivat $mean_{12}(abb)$ mittarilla 99.58 – 112.63% välillä ja $mape_{12}$ -mittarilla 9.66 – 23.86% välissä (Taulukko 5.1). Yksittäisien kuukausien tarkkuudet vaihtelivat abb -mittarilla välillä 58 – 200%. Ennakkovaatimuksena tarkkuudelle oli $abb = 70\%$, joka täytyy keskiarvollisesti mittarilla $mean_{12}(abb)$ kaikissa neljässä kanavassa tarkastelujakson 12 kuukauden aikana. Ensimmäinen kanava täyttää vaatimukset jokaisena kuukautena. Toisessa kanavassa kolme ennustetta saivat vaatimuksen ulkopuolisen tarkkuuden $abb = 64\%, 170\%$

Taulukko 5.1 Kuukausiennustuksien toteutunut tarkkuus ajalta 2015/6 - 2016/5 kaikelle myynninnille ja jaoteltuna neljälle eri kanavalle.

Data	$\text{mean}_{12}(\text{abb})$	mape_{12}
Kaikki	99.24	6.64
Kanava 1	101.63	9.66
Kanava 2	112.85	18.70
Kanava 3	100.14	16.84
Kanava 4	99.58	23.86

ja 200% tarkkailujakson aikana. Kolmannessa kanavassa oli yksi ennuste vaatimusten ulkopuolella arvolla 58% . Neljännessä kanavassa kaksi ennustetta ei täyttänyt vaatimusta arvoilla 59% ja 61%.

Ensimmäisen kanavan ennusteiden tarkkuushajonta pienenee $\text{mean}_{12}(\text{abb})$ -mittarilla ja lähenee täydellistä $\text{mean}_{12}(\text{abb}) = 100\%$ tarkkuutta virityspituutta lisätessä. Mittari mape_{12} saa arvoja 10%:n läheisyydessä virityspituuden vaihtamisesta huolimatta. Koska mittarista riippuen ennustetarkkuus joko paranee tai ei muutu virityspituutta lisättäessä, on perusteltua käyttää kaikkea saatavilla olevaa historiatietoa viritykseen.

Toisen kanavan tarkkuus $\text{mean}_{12}(\text{abb})$ -mittarilla huononee 80-kuukauden jälkeen virityspituutta lisätessä. mape_{12} -mittarin tarkkuus heikkenee noin 70 kuukauden virityspituuden kohdalla mutta saa arvoja 18 – 20%. Näiden kahden tuloksen perusteella mahdollisimman lyhyen, noin 50 – 60 kuukauden, virityspituuden käyttö olisi perusteltua.

Kolmannen kanavan $\text{mean}_{12}(\text{abb})$ -mittarin tarkkuus lisääntyy tasaisesti virityspituutta kasvatettaessa. mape_{12} -mittari puolestaan huononee tasaisesti. Mittarien tarkkuuden trendien ollessa vastakkaiset ei voida antaa yksiselitteistä ohjetta virityspituuden valintaan. Tällöin ennustajan täytyy arvioida kumpi mittari olisi oleellisesti hyödyntäjän kannalta.

Neljännän kanavan kohdalla molemmat mittarit saavat epätarkimmat arvonsa aivan alussa noin 50 kuukauden kohdalla ja 80 kuukauden kohdalla. Käytettäessä kaikkea virityspituutta $\text{mean}_{12}(\text{abb})$ on hyvin lähellä 100% jolloin mape_{12} saa arvoja 24% läheisyydessä. Kaiken historiatiedon käyttö on perusteltua.

Heinäkuulle 2016 saatavia kuukausiennustuksia oli yhteensä 76 kappaletta, joka tar-

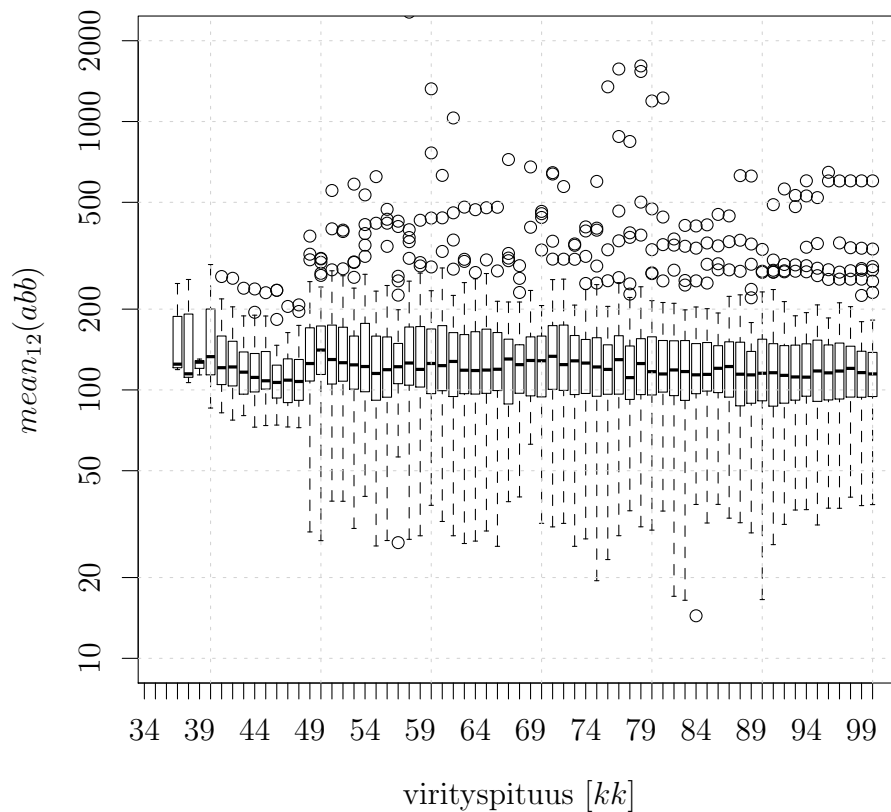
koittaa että 75 % maista sisälsi tuolloin riittävästi mallinnettavaa kuukausittaista myyntihistoriatietoa. Maiden ennustustarkkuuksia pystyi mittaamaan $mean_{12}(abb)$ mittarilla 57 kappaletta ja $mape_{12}$ mittarilla 49 kappaletta (Taulukko 5.2). Maiden kuukausiennustetarkkuudet vaihtelivat suuresti - osa maista on ennustettavissa hyvällä tarkkuudella ja osa ei. Ennustusvirheitä prosenttijohdannaisilla mittareilla tarkasteltaessa pienimmän myynnin maissa on usein suurimmat ennustusvirheet. Ennakkovaatimus oli $abb = 50 - 60\%$, joka toteutui $mean_{12}(abb)$ -mittarilla 41 - 42 maassa (Taulukko 5.2).

Taulukko 5.2 Kuukausiennustuksien toteutunut tarkkuus ajalta 2015/6 - 2016/5 maittain.

$mean_{12}(abb)$	maita	$mape_{12}$	maita
$95 \leq X \leq 105$	6	$0 \leq X \leq 5$	0
$90 \leq X \leq 110$	14	$0 \leq X \leq 10$	0
$80 \leq X \leq 120$	21	$0 \leq X \leq 20$	1
$70 \leq X \leq 130$	33	$0 \leq X \leq 30$	3
$60 \leq X \leq 140$	41	$0 \leq X \leq 40$	12
$50 \leq X \leq 150$	42	$0 \leq X \leq 50$	19
$40 \leq X \leq 160$	45	$0 \leq X \leq 60$	20
$30 \leq X \leq 170$	48	$0 \leq X \leq 70$	20
$20 \leq X \leq 180$	50	$0 \leq X \leq 80$	22
$10 \leq X \leq 190$	51	$0 \leq X \leq 90$	24
$0 \leq X \leq 200$	51	$0 \leq X \leq 100$	28

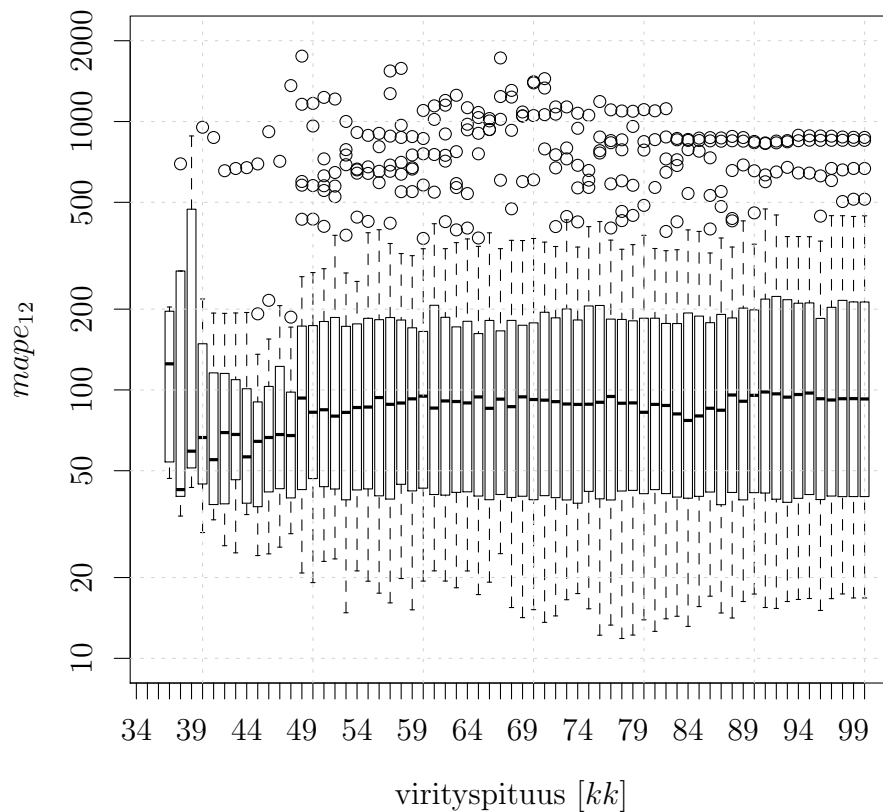
Maiden kuukausiennustetarkkuuksien muutos virityspituuden funktiona on mittarilla $mean_{12}(abb)$ kuvassa 5.1 ja $mape_{12}$ mittarilla kuvassa 5.2. Molemmissa kuvissa näkyy 49 kuukauden virityspituuden kohdalla porrasmainen tarkkuudenmuutos, joka muodostuu usean maan saadessa vasta tuolla virityspituudella 3 vuotta yhtäjaksoista historiatietoa täyteen.

Mittarilla $mean_{12}(abb)$ maiden tarkkuuksien keskiarvot ovat noin 110% läheisyydessä riippumatta virityspituudesta (Kuva 5.1). Virhejakaumien puolikas suurempaan päin ulottuu logaritmisella y-akselilla noin 200% läheisyyteen. Toinen puolikas ulottuu noin 30% läheisyyteen riippuen virityspituudesta. Jakaumien ulkopuoliset virheet alkavat noin 300% läheisyydestä ylöspäin. Tällaisiin suuriin ennustusvirheisiin ei näytä testijaksolla vaikuttavan virityspituus. Katsottaessa kaikkien maiden ennustetarkkuuksien käyttäytymistä muutettaessa virityspituutta ei voida sanoa että virityspituus oleellisesti sitä muuttaisi vaikka näin saattaakin yksittäistapausten kohdalla olla.



Kuva 5.1 Maakohtaisten kuukausiennusteiden tarkkuuden muutos virityspituuden funktiona $mean_{12}(abb)$ -mittarilla.

Mittarilla $mape_{12}$ maiden kuukausiennusteiden virheiden mediaaniarvot ovat lähellä 100% (Kuva 5.2). 49 kuukauden virityspituuden kohdalla oleva porras johtuu myös useamman maan kolmen vuoden historiatiedon vähimmäisvaatimuksen täyttymistä kyseisellä kohdalla. Tarkin kuukausiennustustarkkuusjakaumien neljännes saa arvoja 15 – 40% läheisyydessä riippuen käytetystä virityspituudesta. Kuvasta nähdään että yleisesti ottaen maittain kuukausiennustettaessa virheprosentit ovat $mape_{12}$ -mittarilla suuria ja noin kolme neljästä ennusteesta tuottaa yli 40 prosentin virheen. Huonoin tarkkuusneljännes saa arvoja 200 ja 500 prosentin välistä.



Kuva 5.2 Maakohtaisten kuukausiennusteiden tarkkuuden muutos virityspituuden funktiona $mape_{12}$ -mittarilla.

5.2 Kvartaaliennusteet

Kvartaaliennustuksien tarkkuuksien analysointiaika on kahdeksan kvartaalia 2014/Q2 - 2016/Q1.

Tarkasteluaajan kvarttaaliennustuksien tarkkuus kaikelle myynnille oli $mean_8(abb) = 101.97\%$ ja $mape_8 = 5.49\%$ (Taulukko 5.3). Yksittäisien kvartaalien ennustetarkkuudet tuona aikana olivat $abb = 92 - 115\%$ välillä, jotka kaikki täyttivät ennakkovaatimuksen $abb = 90\%$. Virityspituuden kasvattaminen ei kasvattanut testijaksolla tarkkuutta enää 25 kvartaalin jälkeen. Tällöin tarkkuus asettui $mape_8 = 5 - 6\%$ ja $mean_8(abb) = 102\%$ välille. Kaiken myyntihistorian käyttö on perusteltua vaivatto-

muutensa puolesta vaikka tarkkuuteen sillä ei testijakson aikana ollut vaikutusta.

Taulukko 5.3 Kvartaaliennustuksien toteutunut tarkkuus ajalta 2014/Q2 - 2016/Q1 kaikelle myynninnille ja jaoteltuna neljälle eri kanavalle.

Data	$\text{mean}_8(\text{abb})$	mape_8
All	101.97	5.49
Kanava 1	106.98	11.35
Kanava 2	95.91	15.18
Kanava 3	92.09	23.13
Kanava 4	101.22	4.49

Kanavittain ennustettaessa testijakson tarkkuudet vaihtelivat välillä $\text{mean}_9(\text{abb}) = 92.09 - 106.98\%$ ja $\text{mape}_8 = 4.49 - 23.13\%$ (Taulukko 5.3). Yksittäisien kvartaalien tarkkuudet vaihtelivat kokonaisuudessaan välillä $\text{abb} = 62 - 141$.

Ennakkovaatimusta kanavien kvartaaliennustuksille ei ollut mutta esimerkiksi $\text{abb} = 80\%$ tarkkuuteen kahdeksan kvartaalin vertailuaikana pääsi ensimmäinen kanava jokaisella ennustuksella tarkkuuksilla $\text{abb} = 89 - 130\%$. Toisen kanavan $\text{abb} = 70 - 141\%$ tarkkuuksista yksi ennuste alitti 80% vertailutarkkuuden arvolla $\text{abb} = 70\%$. Kolmannen kanavan $\text{abb} = 62 - 119\%$ tarkkuuksista kaksi ennustetta jäi alle 80% tarkkuuden tuloksilla 62 ja 72. Neljännen kanavan tarkkuudet vaihtelivat välillä $\text{abb} = 89 - 111\%$, joten ne olivat kaikki tarkempia kuin tarkkuusvattimus $\text{abb} = 80\%$.

Ensimmäisen kanavan ennusteet saivat arvoja $\text{mean}_8(\text{abb}) = 102 - 110\%$ välillä virityspituutta muutettaessa. Tarkimmat ennusteet olivat 25 kvartaalin jälkeen tarkkuudella $\text{mean}_8(\text{abb}) = 106 - 108\%$ ja $\text{mape}_8 = 11 - 12\%$. Testijakson aikana kaiken myyntihistorian käyttö ennustukseen on perusteltua.

Toisen kanavan ennustetarkkuudet saivat arvoja $\text{mean}_8(\text{abb}) = 91.3 - 104.1\%$ testijaksolle virityspituutta muutettaessa. Tarkimmat ennustukset olivat noin 25 kvartaalin jälkeen tarkkuuksilla $\text{mean}_8(\text{abb}) = 96\%$ ja $\text{mape}_8 = 14.8$. Ennusteet tarkentuivat virityspituutta lisätessä, joten kaiken myyntihistorian käyttö on suositeltavaa.

Kolmannen kanava ennustetarkkuudet saivat arvoja $\text{mean}_8(\text{abb}) = 95 - 178\%$ eri virityspituuksilla. Tarkimmat ennusteet syntyivät yli 28 kvartaalin myyntihistorian käytöllä mittarilla $\text{mean}_8(\text{abb})$. Tällöin ennustetarkkuudet olivat noin 95% . mape_8 mittarilla virheet nousivat $16 - 23$ tasaisesti historiaa lisätessä. Suositeltava virityshistorian pituus riippu käytettävästä tarkkuusmittarista eikä ole yksiselitteinen.

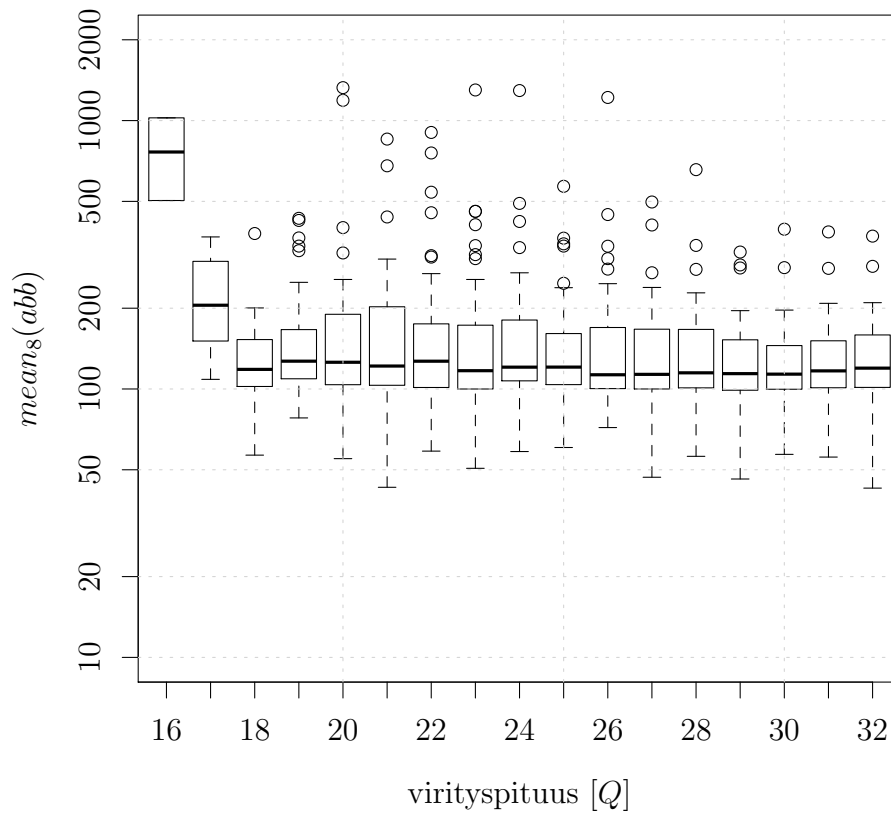
Neljannen kanavan ennustetarkkuudet saivat arvoja $mean_8(abb) = 94.2 - 102.8$. Mo-lemmilla mittareilla 25 kvartaalin jälkeen lisätty virityspituus tarkensi ennusteita. Kaikella historiatiedolla $mean_8$ päättyi 101.2% tarkkuuteen ja $mape_8$ neljään prosenttiin. Kaiken myyntihistorian käyttö oli perusteltua, jotta testidatalla sai tarkimmat tulokset.

Kvartaalille 2016 Q3 saatavia maakohtaisia ennusteita oli yhteensä 80 kappaletta, joka tarkoittaa että 79% maista sisälsi riittävästi mallinnettavaa kvartaalittaista myyntihistoriaa. Maiden ennustarkkuutta pystyi arvioimaan $mean_{12}(abb)$ mittarilla 64 kappaletta ja $mape_8$ mittarilla 66 kappaletta (Taulukko 5.4).

Taulukko 5.4 Kvartaaliennustuksien toteutunut tarkkuus ajalta 2014/Q2 - 2016/Q1 maittain.

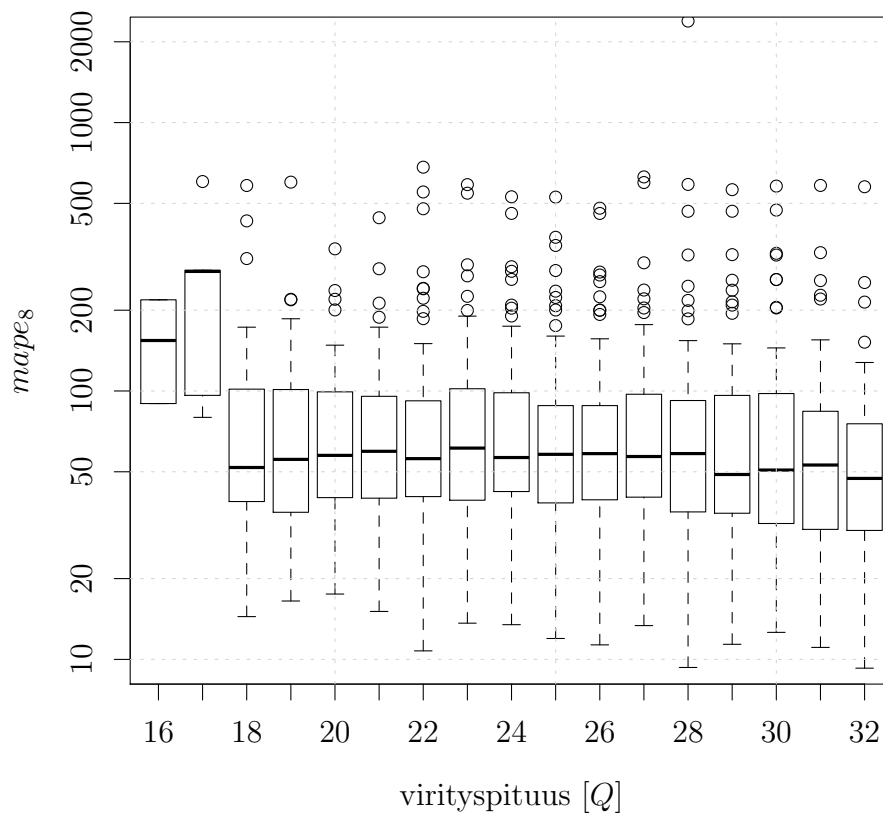
$mean_{12}(abb)$	maita	$mape_{12}$	maita
$95 \leq X \leq 105$	6	$0 \leq X \leq 5$	0
$90 \leq X \leq 110$	14	$0 \leq X \leq 10$	0
$80 \leq X \leq 120$	21	$0 \leq X \leq 20$	1
$70 \leq X \leq 130$	33	$0 \leq X \leq 30$	3
$60 \leq X \leq 140$	41	$0 \leq X \leq 40$	12
$50 \leq X \leq 150$	42	$0 \leq X \leq 50$	19
$40 \leq X \leq 160$	45	$0 \leq X \leq 60$	20
$30 \leq X \leq 170$	48	$0 \leq X \leq 70$	20
$20 \leq X \leq 180$	50	$0 \leq X \leq 80$	22
$10 \leq X \leq 190$	51	$0 \leq X \leq 90$	24
$0 \leq X \leq 200$	51	$0 \leq X \leq 100$	28

Maiden kvartaaliennustusten tarkkuudella ei ollut ennakkovaatimusta. Esimerkiksi $mean_8(abb) = 60 - 70\%$ tarkkuuteen ylsi 33 – 41 maata (Taulukko 5.4). Kaikkien maiden kvartaaliennustuksia tarkastellessa ennustustarkkuus asettuu lopulliseen tarkkuuteensa molemmilla mittareilla noin 18 kvartaalin virityspituuden jälkeen (Kuvat 5.4 ja 5.3).



Kuva 5.3 Maakohtaisten kvartaaliennusteiden tarkkuuden muutos virityspituuden funktiona $mean_8(abb)$ -mittarilla.

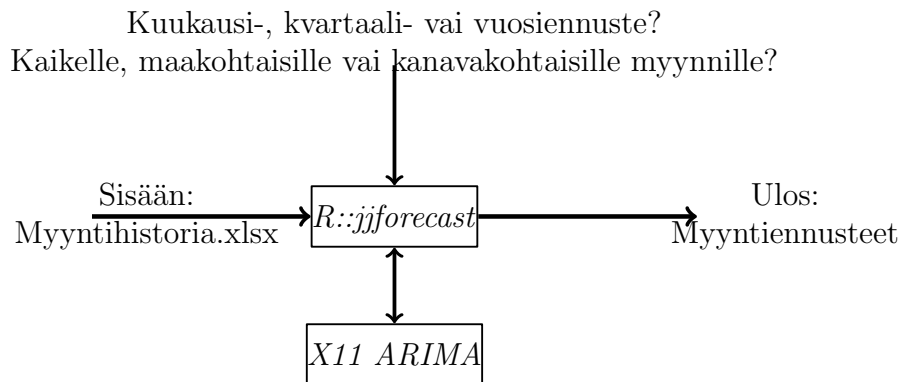
Kuukausiennustuksien tapaan maittain ennustettaessa on suuria tarkkuuseroja eri maiden väleillä ja jokaisen maan soveltuminen ennustamiseen täytyy tarkastella erikseen. Pisimmällä 32 kvartaalin virityspituudella kaikista maista puolet ovat tarkkuuden $mean_8(abb) = 100 - 180\%$ ja $mape_8 = 40 - 80\%$ sisällä. Tarkin neljännes saa tuolloin ennustuksia tarkkuudella $mape_8 = 9 - 40\%$ (Kuva 5.4).



Kuva 5.4 Maakohtaisten kvartaaliennusteiden tarkkuuden muutos virityspituuden funktiona $mape_8$ -mittarilla.

6. TULOKSET

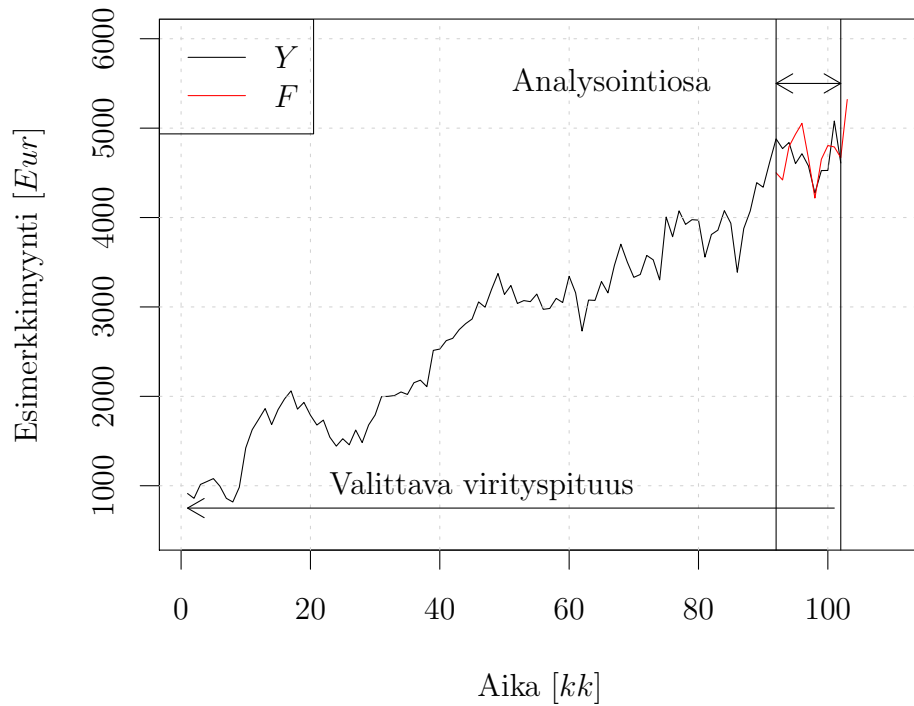
Työssä tutkittiin miten tilastollisten ennusteiden tarkkuutta voidaan mitata ja miten ennustemallin viritykseen käytetty historianpituus vaikuttaa saataviin tarkkuuksiin. Tutkimusta varten päädyttiin virittämään Arima x11-ennustemalli tuottamaan ennusteita (Kuva 6.1). Malli jakaa annetun myyntihistorian kausi-, trendi- ja satunnaiskomponentteihin, joiden avulla ennusteet muodostetaan.



Kuva 6.1 Ennustusjärjestelmän toiminnallinen rakenne. Sisääntulona toimii annettu *.xlsx -tiedosto, josta lasketaan valitut ennusteet näytölle. Taustalla ennustuskoneena käytetään X11-Arima -ennustusohjelmaa. Käyttäjä valitsee halutun ennustusjakson sekä kohteen. jfforecast toimii tulkkina käyttäjän ja ennustusohjelman välillä hoitaen datankäsittelyn.

Tarkkuuden tilastollista mittaamista varten päädyttiin 12 kuukauden analysointijaksoon kuukausiennusteille ja 8 kvartaalin analysointijaksoon kvartaaliennusteille (Kuva 6.2). Kyseisten jaksojen aikana tutkittiin ennustetun sarjan ja todellisen myyntihistorian välistä eroa. Jokaisen ennusteen tuottamiseen käytettiin vain kyseisellä hetkellä saatavissa olevaa historiatietoa. Ennustetarkkuuden muutosta virityspituuden funktiona tarkasteltiin ennustamalla käyttäen kaikkia mahdollisia virityspituuksia (Kuva 6.2) ja tarkastelemalla analysointijakson aikaisia ennustevirheitä.

Ennustesarjan ja todellisen myynnin välisen virheen mittaamiseen on käytettävissä



Kuva 6.2 *Analysointiosan ennustuksien tarkkuuden muuttumista virityspituuden funktiona voidaan tarkastella laskemalla ennustukset eri virityspituuksilla. Analysointiosan pituus on kuukausiennustuksille 12 kuukautta ja kvartaaliennustuksille 8 kvartaalia.*

monia mittareita. Jokainen mittari näyttää vain yhden näkökulman virheen luonteeseen, joten päädyttiin käyttämään useampaa mittaria paremman yleiskuvan luomiseksi. Yrityksen prosenttijohdannaisen mittarin (Kaava 6.1) käyttö helpottaa tulosten vertailua sisäisesti.

$$ABB_t = \frac{\text{Todellinen myynti}}{\text{Ennustettu myynti}} \cdot 100 = \frac{100Y_t}{F_t} \quad (6.1)$$

Mittari $MABB_n$ (Kaava 6.2) on mittarin ABB_t tuottamien tarkkuuksien keskiarvo analysointijakson aikana. Näin tarkkuus voidaan ilmaista yhtenä lukuna, jota voidaan verrata muihin samalla tavalla saatuihin tarkkuuksiin.

$$MABB_n = \frac{1}{n} \sum_{t=1}^n ABB_t = \frac{100}{n} \sum_{t=1}^n \frac{Y_t}{F_t} \quad (6.2)$$

Akateemisena mittarina käytettiin yhtä eniten käytetyistä mittareista, absoluuttista prosentuaalista virhettä $MAPE_n$ (Kaava 6.3), jota käytetään analysointijakson aikana

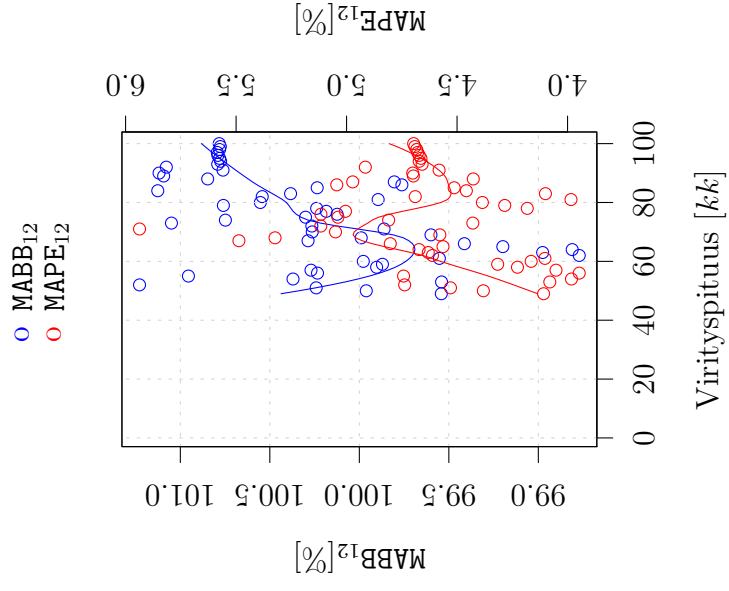
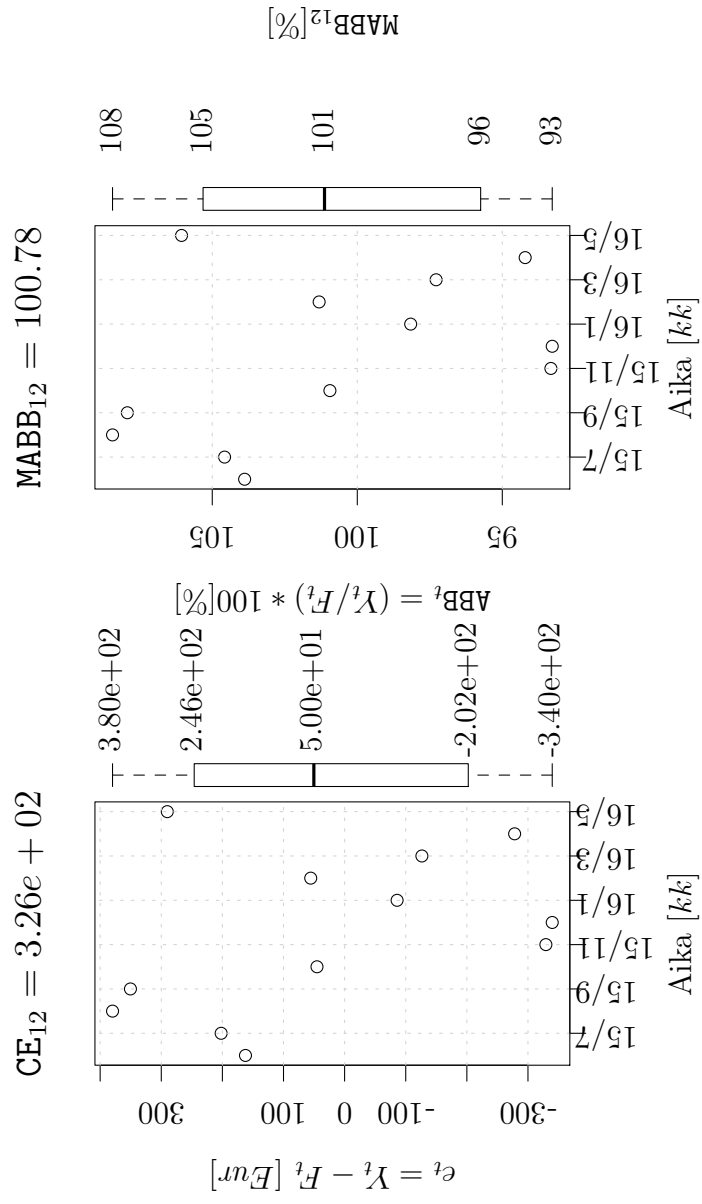
tapahtuvien virheiden ilmaisemiseen yhdellä luvulla.

$$\text{MAPE}_n = \frac{100}{n} \sum_{t=1}^n \frac{|Y_t - F_t|}{Y_t} \quad (6.3)$$

Euromääräinen virhe e_t , sekä sen kumulatiivinen summa kertovat kuinka merkittävä ennustevirhe on esimerkiksi budjetoinnissa ja kumoavatko virheet toisensa analysointijakson aikana.

$$\text{CE}_n = \sum_{t=1}^n e_t = Y_t - F_t \quad (6.4)$$

Valittuja numeerisia mittareita käyttäen on muodostettu kolme tarkkusgraafia esittämään ennustetarkkuus. Esimerkin satunnaiskululla muodostetulle myyntihistorialle (Kuva 6.2) on muodostettu ennustetarkkuus samalla tavalla kuin kaikille ennustetaville myyntihistorioille (Kuva 6.3). Vasemman puoleiseen graafiin on sijoitettu euromääräiset virheet analysointijakson aikana. Keskimmäinen graafi kuvaa yksittäiset ennustetarkkuudet ja niiden keskiarvon yrityksen sisäisellä tarkkuusmittarilla. Oikeanpuoleinen graafi kuvaa tarkkuuden muuttumista virityspituuden funktiona MAPE_n ja MABB_n mittareiden avulla. Satunnaiskulun ennustamisen tarkkuudesta ja virheen suuruudesta näkee järjestelmän yleisen suoritustason. Satunnaiskulun ennustustarkkuus on valitusta virityspituudesta riippuen $\text{MAPE}_{12} = 4 - 6\%$ ja $\text{MABB}_{12} = 98.5 - 101.5\%$. ABB_t -mittari sai arvoja $93 - 108\%$ välillä 12 kuukauden analysointijakson aikana.



Kuva 6.3 Satunnaiskullalla muodostetun myynnin kuukausiennustuksen tarkkuus.

Tällä tavoin suoritettut ennustamiset ja mittaukset täyttivät yrityksen ennakkovaatimukset (Taulukko 6.1) koko myynnin ennustamisessa kaikissa tapauksissa, kanavittain tehdyissä ennusteissa 88% täytti odotukset ja maittain tehtävistä kuukausiennusteista täyttivät 55% ennako-odotukset (Taulukko 6.2).

Taulukko 6.1 *Ennustusjärjestelmälle asetetut tarkkuusvaatimukset ABB_t -mittarilla (Kaava 6.1). Sulkeissa olevat tarkkuudet ovat lisätty analyysia varten eivätkä siten ole virallisia vaatimuksia.*

Data			
Kaikki	Kanavat	Maat	Ennusteaika
80%	70%	50-60%	Kuukausi
90%	[80%]	[60-70%]	Kvartaali
-	-	-	Loppuvuosi

Taulukko 6.2 *Yksittäisien ennustusten suoriutuminen tutkimuksessa esitellyllä menetelmällä mitattuna ABB_t -mittarilla (Kaava 6.1) suhteessa annettuihin tarkkuusvaatimuksiin (Taulukko 4.1). Analysointiaika tarkkuuden mittaamista varten on kuukausiennusteille vuosi ja kvartaaliennusteille kaksi vuotta. Sulkeissa oleviin suoriutumisprosentteihin ei ollut saatavissa lähtövaatimusta, joten arvot ovat laskettu kuvitteellisista vaatimuksista.*

Data			
Kaikki	Kanavat	Maat	Ennusteaika
100%	88%	55%	Kuukausi
100%	[91%]	[41%]	Kvartaali
-	-	-	Loppuvuosi

Virityspituuden valinnasta ei voida saatujen tuloksien perusteella antaa yksiselitteistä ohjetta käytetylle datalle. Saatavat tarkkuudet eivät käytetyillä mittareilla kokonaisuudessaan parane tai heikkene virityspituutta muutettaessa. Sen sijaan sarjoja yksittäin tutkittaessa osan ennusteet tarkentuivat, osan heikentyivät ja osaan ei vaikuttanut ollenkaan virityspituuden muutokset. Osaan sarjoista käytetyt mittarit näyttivät ristikkäisiä tuloksia.

7. YHTEENVETO

Työn tarkoituksena oli sovittaa Arima x11 -malli myynnin ennustamiseen ja tutkia kuinka tarkkoihin ennusteisiin sen oletusasetuksilla avulla päästään. Tutkittavat ennusteet olivat kuukausi-, kvartaali- ja loppuvuosiennusteet kaikelle, kanavittaisille ja yksittäisille maille. Kuukausittain tehtävien loppuvuosiennusteiden tarkkuutta ei mitattu eikä analysoitu suuren hajonnan vuoksi. Mallin ennustetarkkuuksia verrattiin arvioituihin ennustetarkkuuksiin usealla eri tarkkuusmittarilla. Lisäksi tutkittiin mallin viritykseen käytetyn historiatiedon pituuden vaikutusta saataviin ennustetarkkuuksiin.

Ennustetarkkuuden mittaamista varten rakennettiin erillinen simulointijärjestelmä, joka simuloi ennustusjärjestelmän tuottamia ennusteita ja mittaa virhettä määritellyin mittarein. Ennustamisessa ja sen tarkkuuden mittaamisessa tehtävien valintojen vuoksi ei aina ole yksiselitteistä kuvastaako saatava ennustetarkkuus mallin, mallintajan vai mittaajan etevyyttä. Arvioitaessa mallin sopivuutta suunniteltuun käyttötilanteeseen täytyy kiinnittää huomiota tarkkuusmittarin valintaan sillä yhden parantuessa toinen saattaa huonontua.

Tässä tutkimuksessa esitetyn simulointitavan perusteella Arima x11 - malli täytti annetut vaatimukset taulukon 7.1 mukaisesti. Kuukausi ja kvartaaliennusteet kai-

Taulukko 7.1 Yksittäisien ennustusten suoriutuminen tutkimuksessa esitellyllä menetelmällä mitattuna ABB_t -mittarilla (Kaava 4.1) suhteessa annettuihin tarkkuusvaatimuksiin (Taulukko 4.1). Analysointiaika tarkkuuden mittaamista varten on kuukausiennusteille vuosi ja kvartaaliennusteille kaksi vuotta. Sulkeissa oleviin suoriutumisprosentteihin ei ollut saatavissa lähtövaatimusta, joten arvot ovat laskettu kuvitteellisista vaatimuksista.

Data			
Kaikki	Kanavat	Maat	Ennusteaika
100%	88%	55%	Kuukausi
100%	[91%]	[41%]	Kvartaali
-	-	-	Loppuvuosi

kelle myynnille täyttivät ennakko-odotukset koko analysointijakson ajan. Pienenettäessä ennustusalueita myös tarkkuudet ja suoriutuminen ennakko-osotuksiin nähden huonontui.

Numeeriseen tarkkuuslukemaan vaikuttavat ainakin valittu matemaattinen malli, sen sovittajan aihealueen osaaminen, tarkkuusmittarin valinta, tarkkuusmittauksien suorittaminen sekä itse ennustettavan datan säännöllisyys. Näistä syistä johtuen tämän työn tarkkuusmittarien lukemista ei voida tehdä suoria johtopäätöksiä käytetyn mallin hyvyydestä verrattuna toisissa tutkimuksissa käytettyihin malleihin ja niiden tarkkuuksiin. Jotta mallien sopivuutta käyttötilanteeseen voitaisiin verrata, täytyisi ne virittää yksitellen samalle datalle ja mitata niiden suoriutuminen samalla tavoin.

Prosenttijohdannaisilla MAPE- ja MABB- mittareilla suurimmat ennustevirheet syntyivät pienimpien ja vaihtelevimpien myyntimäärien ennustamisessa. Suurempaa kokonaisuutta ennustettaessa myös prosentuaaliset myyntimäärän muutokset ovat usein pienempiä, joka voi olla osasy syy käytetyn mallin niissä paremmin suoriutumiseen.

Euromääräisen virheen kumulatiivisen summan avulla pystyi arvioimaan MAPE- ja MABB- mittarien ennustusvirheen merkitsevyyttä ja kumoavatko virheet euromääräisesti toisensa analysointiaikana. Ennustettavissa sarjoissa ei kokonaisuutena ollut selvää ennusteiden yli- tai alisuoriutumista analysointiajalla, joten tilanne täytyy tarkastaa sarjakohtaisesti simulointituloksista.

Mikäli historiallinen data ei kuvaa enää nykyistä toimialaa riittävästi, voi olla että pidemmällä historiatiedolla ei saavuteta parempia ennustetuloksia. Virityspituuden valinnan vaikutusta tarkkuuteen mitattiin MAPE- ja MABB mittareilla. Vaikka kokonaisuutena selvää sääntöä virityspituuden valinnalla tarkkuuteen ei havaittu, huomattiin sillä olevan suuri vaikutus joidenkin yksittäisten sarjojen kohdalla. Ennustetulokset saattoivat parantua, heikentyä tai ajalehtia virityspituutta lisätessä. Lisäksi mittareista toinen saattoi tarkentua toisen heikentyessä virityspituutta muutettaessa. Käytetyn mittarin valinnan tärkeys korostuu mikäli ennustemallin tarkkuutta yritettäisiin parantaa jatkokehityksellä. Tämä on osa laajempaa kokonaisuutta mielekkäiden mittarien valinnasta.

Trendin käännekohdan ennustaminen on vaikeaa, kuten myös mallin suoriutumisen mittaaminen sen aikana. Malli saattaa siten toimia hyvin keskimäärin mutta ei tren-

din suunnan muuttuessa muutosajankohtien ollessa usein vähemmistönä tilastollisessa laskennassa, jota puolestaan käytetään monesti mallinnuksessa ja tarkkuuden mittauksessa.

Useamman kuin yhden mittarin käytöllä huomattiin saavutettavan parempi näkemys virheen luonteesta ja ennusteiden käytettävyydestä suunniteltuun tilanteeseen. Ennusteiden loppukäyttäjien tietotaito vaikuttaa myös osaltaan käytettävien mittarien valintaan - yksinkertainen mittari on helpompi ymmärtää kuin monimutkainen, joka helpottaa ennusteiden satunnaisempaa käyttäjää.

Matemaattiset mallit toimivat eri tavoin ja siksi niillä on eri vahvuusalueita. Käytetyllä Arima x11-mallilla on 3 vuoden yhtäjaksoisen historiatietovaatimus ennen kuin se pystyy ennustamaan, josta seuraa että uutta ennustettavaa kokonaisuutta, kuten uutta maata, ei voida ennustaa heti alusta lähtien. Käytettävimmän mallin löytämiseksi pitäisikin etsiä sellaista, joka toimii niillä osa-alueilla parhaiten kuin on oleellista liiketoiminnan kannalta.

Yrityksen toimintaympäristössä tutkimuksen tuloksena kaiken myynnin kuukausi- ja kvartaaliennusteiden tarkkuus on riittävä ja toimii siten apuna niitä laadittaessa. Neljän kanavan ennusteetkaan eivät jääneet kauksi tavoitteista, joten niitäkin voidaan käyttää apuna ennusteiden laadinnassa. Maittain tehtävien ennusteiden tarkkuudet vaihtelivat suuresti maiden välillä. Tästä johtuen niitä käytettäessä pitää simulointituloksista arvioida mahdollinen ennustetarkkuus ja onko se soveltuva käyttötarkoitukseen. Yleisesti ottaen käytetty ennustejärjestelmä toimii hyvin kausivaihteluiden mallintamisessa mutta ei trendin ennustamisessa. Tätä voidaan käyttää hyödyksi ja tarkastaa onko jokin tapahtunut muutos normaalia vaihtelua vai onko kyseessä jo trendin suunnan muutos.

Jatkotutkimuksena voidaan virittää useampi malli rinnakkain vertailtavaksi. Yleisen käytettävyydsmittarin, joka koostuisi painotetuista tarkkuusmittareista, muodostamista ja käyttämisestä voitaisiin tutkia. Lisäksi voitaisiin tutkia parantaisiko historiatiedon inflaatiokorjaaminen nykyisen mallin ennustetarkkuuksia.

Anti ennustajien hyödyntäjille on että kannattaa kysyä oletuksia millä ennusteet ovat laskettu ja onko saatavilla useampia mittareita mallin suoritusmittareiksi. Voi myös selvittää perustuuko ennuste pelkkään tekniseen data-analyysiin ja syy- ja seuraussuhteiden tulkintaan vai onko taustalla myös oikeaa tietoa miten on päätetty toimia, joka on aina ennustamista tarkempaa tietoa.

Aikasarjan mallinnuksessa kannattaa kiinnittää huomiota mallin tarkkuusmittarin valintaan ja mallin toiminta-alueanalyysiin, jotta mallin heikkoudet ja vahvuudet voidaan havaita.

LÄHTEET

- [1] X. Zhao, J. Xie, and R. Lau, “Improving the supply chain performance: Use of forecasting models versus early order commitments,” vol. 39, pp. 3923–3939, *International Journal of Production Research*, 2001. Saatavissa: <http://dx.doi.org/10.1080/00207540110072236>.
- [2] M. Babai, M. Ali, J. Boylan, and A. Syntetos, “Forecasting and inventory performance in a two-stage supply chain with arima(0,1,1) demand: Theory and empirical analysis,” DeBEM-Bordeaux Management School, 680 cours de la Libération, 33405 Talence Cedex, France and King Saud University, Saudi Arabia and Bucks New University, United Kingdom and University of Salford, United Kingdom, Science Direct, 2011. Saatavissa: <http://www.sciencedirect.com/science/article/pii/S0925527311003902>.
- [3] P. Danese and M. Kalchschmidt, “The role of the forecasting process in improving forecast accuracy and operational performance,” vol. 131, pp. 204 – 214, *Science direct*, 2011.
- [4] L. Ljung, “Modelling of dynamic systems,” Prentice Hall Information and System sciences, pp. 14, 261, PTR Prentice Hall, 1994.
- [5] R. Hyndman, “Forecasting: principles and practice.” Available: <http://otexts.org/fpp/>, 2017.
- [6] R. Yaffee and M. McGee, “Introduction to time series analysis and forecasting with applications of sas and spss,” No. 0-12-767870-0, (Harcourt Place, 32 Jamestown Road, London NW1 7BY, UK), pp. 5,45, 52, Academic Press, Inc, 2000.
- [7] P. Saikkonen, “Stationaariset aikasarjat.” Helsingin Yliopisto. Matematiikan ja tilastotieteenlaitos. Luentomoniste [online], 2015. Available: https://wiki.helsinki.fi/download/attachments/164335189/StatAika10_3.pdf.
- [8] A. Luoma, “Aikasarja-analyysi i.” Tampereen Yliopisto. Kurssimateriaali [online], 2015. Available: <http://people.uta.fi/~al18853/luentod.pdf>.
- [9] J. DeCoster, “Applied linear regression notes 1,” (348 Gordon Palmer Hall, Box 870348, Tuscaloosa AL 35487-0348), pp. 5,9,10, Department of Psychology

- University of Alabama, 2007. Saatavissa: <http://www.stat-help.com/cohen%202007-11-13.pdf>.
- [10] T. A. Runkler, “Data analytics - models and algorithms for intelligent data analysis,” pp. 56,57, Springer Vieweg, 2012.
 - [11] A. J. Jaffe and H. F. Spirer, “Misused statistics - straight talk for twisted numbers,” pp. 13–22,46, Marcel Decker, Inc, 1987.
 - [12] J. Everette S. Gardner, “Exponential smoothing: The state of the art,” vol. 4 of *Journal of Forecasting*, pp. 1–28, Wiley, 1985. Saatavissa: <http://onlinelibrary.wiley.com/doi/10.1002/for.3980040103/abstract>.
 - [13] J. Everette S. Gardner, “Exponential smoothing: The state of the art — part ii,” vol. 22 of *International Journal of Forecasting*, pp. 637 – 666, Elsevier, 2006. Saatavissa: <https://www.sciencedirect.com/science/article/pii/S0169207006000392>.
 - [14] E. G. Jr and E. McKenzie, “Why the damped trend works,” vol. 62 of *The Journal of the Operational Research Society*, pp. 1177–1180, Palgrave Macmillan Journals on behalf of the Operational Research Society, 2011. Saatavissa: <http://www.jstor.org/stable/20868958>.
 - [15] F. Chen, J. K. Ryan, and D. Simchi-Levi, “The impact of exponential smoothing forecasts on the bullwhip effect,” vol. 47, pp. 269–286, Naval Research Logistics, 2000. Saatavissa: [http://onlinelibrary.wiley.com/doi/10.1002/\(SICI\)1520-6750\(200006\)47:4%3C269::AID-NAV1%3E3.0.CO;2-Q/abstract](http://onlinelibrary.wiley.com/doi/10.1002/(SICI)1520-6750(200006)47:4%3C269::AID-NAV1%3E3.0.CO;2-Q/abstract).
 - [16] X. Wang and S. M. Disney, “The bullwhip effect: Progress, trends and directions,” vol. 250, pp. 691 – 701, Elsevier B.V., 2016. Saatavissa: <https://www.sciencedirect.com/science/article/pii/S0377221715006554>.
 - [17] R. Metters, “Quantifying the bullwhip effect in supply chains,” vol. 15 of *Journal of Operations Management*, pp. 89–100, 1997. Saatavissa: <http://www.sciencedirect.com/science/article/pii/S0272696396000988>.
 - [18] T. Kohonen, “The self organized maps (som),” CIS - Laboratory of computer and information science - Adaptive informatics research center, 1996. Saatavissa: <http://www.cis.hut.fi/projects/somtoolbox/theory/somalgorithm.shtml>.

- [19] H. Turner and D. Firth, *Generalized nonlinear models in R: An overview of the gnm package*, 2015. R package version 1.0-8.
- [20] F. Baty, C. Ritz, S. Charles, M. Brutsche, J.-P. Flandrois, and M.-L. Delignette-Muller, “A toolbox for nonlinear regression in R: The package nlstools,” *Journal of Statistical Software*, vol. 66, no. 5, pp. 1–21, 2015.
- [21] C. Chu and G. Zhang, “A comparative study of linear and nonlinear models for aggregate retail sales forecasting,” *International journal of Production Economics*, 2003. Saatavissa: <http://www.sciencedirect.com/science/article/pii/S0925527303000689>.
- [22] D. Stojanović, “Modelling sales,” vol. 45 of *The Journal of the Operational Research Society*, pp. 1060–1067, Palgrave Macmillan Journals on behalf of the Operational Research Society, 1994. Saatavissa: <http://www.jstor.org/stable/2584147>.
- [23] J. Layer and T. O’Brien, “Sales forecasting using cyclical analysis,” p. 26, *Journal of Business and Industrial Marketing*, 1988. Saatavissa: <http://dx.doi.org/10.1108/eb006048>.
- [24] P. Koellinger and A. Thurik, “Entrepreneurship and the business cycle,” *The MIT Press Journals*, 2012. Saatavissa: http://www.mitpressjournals.org/doi/pdf/10.1162/REST_a_00224.
- [25] S. Schmitt-Grohe and M. Uribe, “What’s news in business cycles,” vol. 80, pp. 277 – 2764, *Econometrica*, 2012. Saatavissa: <http://www.jstor.org/stable/23357239>.
- [26] M. Camacho and J. Martinez-Martin, “Monitoring the world business cycle,” *Science Direct*, 2015. Saatavissa: <http://www.sciencedirect.com/science/article/pii/S0264999315002606>.
- [27] M. Halling, J. Yu, and J. Zechner, “Leverage dynamics over the business cycle,” No. 122 in *Journal of Financial Economics*, pp. 21–41, *Science Direct*, 2016. Saatavissa: <http://www.sciencedirect.com/science/article/pii/S0304405X16301222>.
- [28] C. Sax, *seasonal: R Interface to X-13-ARIMA-SEATS*, 2017. R package version 1.6.1. Saatavissa: <https://CRAN.R-project.org/package=seasonal>.

- [29] P. Cichosz, “Data mining algorithms explained using r,” (John Wiley and Sons, Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom), p. 8, Warsaw University of Technology, Department of Electronics and Information Technology, Wiley, 2015.
- [30] N. Zumel and J. Mount, “Practical data science with r,” No. ISBN 9781617291562, (Manning Publications Co.,20 Baldwin Road,PO Box 261,Shelter Island, NY 11964), p. 109, Manning, 2014.
- [31] B. Lanz, “Machine learning with r,” No. ISBN 978-1-78216-214-8, (Packt Publishing Ltd.,Livery Place,35 Livery Street,Birmingham B3 2PB, UK.), p. 16, Packt Publishing Ltd, 2013.
- [32] J. G. D. Gooijer and R. J. Hyndman, “25 years of time series forecasting,” No. 22 in *International Journal of Forecasting*, pp. 16, 443–473, Science Direct, 2006. Saatavissa: <https://doi.org/10.1016/j.ijforecast.2006.01.001>.
- [33] “Tshernobyl 16 vuotta myöhemmin,” in *ATS-INFO, Huhtikuu 2002*, Suomen Atomiteknillinen Seura. Saatavissa: https://www.ats-fns.fi/images/files/general/ATS_info_tshernobyl16v_2002.pdf.
- [34] “Vuonna 1997 kohistiin klooni­lammas dollysta,” in *Tiede 35 Vuotta*, Sanoma. Saatavissa: http://www.tiede.fi/artikkeli/uutiset/vuonna_1997_kohistiin_klooni­lammas_dollysta.
- [35] M. P. Clements, “Real-time factor model forecasting and the effects of instability,” *Science Direct*, 2014. Saatavissa: <http://www.sciencedirect.com/science/article/pii/S0167947315000225>.
- [36] R. Rieg, “Do forecasts improve over time? - a case study of the accuracy of sales forecasting at a german car manufacturer,” vol. 18 of *International Journal of Accounting and Information Management*, pp. 220–236, Emerald Insight, 2010. Saatavissa: <http://dx.doi.org/10.1108/18347641011068974>.
- [37] C. Bergmeir, M. Constantini, and J. Benitez, “On the usefulness of cross-validation for directional forecast evaluation,” pp. 132 – 143, Department of Computer Science and Artificial Intelligence, E.T.S. de Ingenierías Informática y de Telecomunicación, University of Granada, CITIC, iMUDS, Spain and Department of Economics and Finance, Brunel University, United Kingdom, *Science Direct*, 2014. Saatavissa: <http://www.sciencedirect.com/science/article/pii/S0167947314000310>.

- [38] C. Bergmeir and J. Benitez, “On the use of cross-validation for time series predictor evaluation,” vol. 192, pp. 192 – 213, Department of Computer Science and Artificial Intelligence, E.T.S. de Ingenierías Informática y de Telecomunicación, University of Granada, CITIC, iMUDS, Spain, Science Direct, 2012. Saataavissa: <http://www.sciencedirect.com/science/article/pii/S0020025511006773>.
- [39] S. Makridakis, A. Andersen, R. Carbone, R. Fildes, M. Hibon, R. Lewandowski, J. Newton, E. Parzen, and R. Winkler, “The accuracy of extrapolation (time series) methods: Results of a forecasting competition,” vol. 1 of *Journal of Forecasting*, pp. 111–153, Research Gate, 1982. Saataavissa: https://www.researchgate.net/publication/245162829_Accuracy_of_Forecasting_An_Empirical_Investigation.
- [40] H. L. Linder and J. K. Horne, “Evaluating statistical models to measure environmental change: A tidal turbine case study,” vol. 84 of *Ecological Indicators*, p. 765–792, Elsevier, 2018. Saataavissa: <https://www.sciencedirect.com/science/article/pii/S1470160X1730609X>.
- [41] D. J. Murray-Smith, *Testing and Validation of Computer Simulation Models - Principles, Methods and Applications*. Springer International Publishing AG, 2015.
- [42] J. McKenzie, “Mean absolute percentage error and bias in economic forecasting,” vol. 113, pp. 259–262, School of Economics, Merewether Building H04, University of Sydney, Sydney 2006, Australia, December 2011. Saataavissa: <http://www.sciencedirect.com/science/article/pii/S0165176511003119>.
- [43] V. Telkki, “Fysiikan laboratoriotyöt 1.” material at noppa.oulu.fi, [online], 2016. Available: https://noppa.oulu.fi/noppa/kurssi/761121p/materiaali/761121P_luentomoniste.pdf.
- [44] “X-13arima-seats seasonal adjustment program,” United States Census Bureau, 2017. Saataavissa: <https://www.census.gov/srd/www/x13as/>.
- [45] “The r project for statistical computing,” The R Foundation, 2017. Saataavissa: <https://www.r-project.org/>.